Transcript of June 19, 2001 Meeting

Please Note: This transcript has not been edited and CMS makes no representation regarding its accuracy.

00001

HEALTH CARE FINANCING ADMINISTRATION
Medicare Coverage Advisory Committee
Meeting of the Diagnostic Imaging Panel

June 19, 2001

Baltimore Convention Center
One West Pratt Street
Baltimore, Maryland

00002

Panelists

Chairperson
Frank J. Papatheofanis, MD, PhD, MPH
Vice-Chairperson
Barbara J. McNeil, MD, PhD

Voting Members
Carole R. Flamm, MD, MPH

Jeffrey C. Lerner, PhD
Michael Manyak, MD
Donna C. Novak, ASA, MAAA, MBA
Steven Guyton, MD

Industry Representative
Michael S. Klein, MBA

Guests
Arnold J. Krubsack, PhD, MD
Jeff Abrams, MD

HCFA Liaison
Sean R. Tunis, MD, MSc
Executive Secretary
Janet Anderson

00003

TABLE OF CONTENTS

	Page
Opening Remarks	
Janet Anderson	5
Sean R. Tunis, MD, MSc	7
Charge to the Panel	
Frank J. Papatheofanis, MD, PhD, MPH	9
Presentation of questions for FDG-PET and breast cancer	
Mitchell Burken, MD	10
Presentation of technology assessment	
David Samson	13
Scheduled Public Comments	
Sanjiv Sam Gambhir, MD, PhD	78
F. David Rollo, MD, PhD, FACC, FACNP	97
Steven Larson, MD	120
Lunch	132

00004

TABLE OF CONTENTS (Continued)

Open Public Comments	134
Open Panel Deliberation	151
Further Discussion and Final Panel Recommendations	201
Closing remarks	289
Adjournment	291

00005

PANEL PROCEEDINGS

(The meeting was called to order at 8:35 a.m., Tuesday, June 19, 2001.)

MS. ANDERSON: Good morning and welcome, committee chairperson, members and guests. I am Janet Anderson, executive secretary of the Diagnostic Imaging Panel of the Medicare Coverage Advisory Committee. The committee is here today to hear and discuss presentations regarding the diagnosing and staging of breast cancer using positron emission tomography scanning technology.

In evaluating the evidence presented to you today, HCFA encourages the panel to consider

all relevant forms of information, including but not limited to professional society statements, clinical guidelines, and other testimony you may hear during the course of this panel meeting.

The following is for the record: For today's panel meeting, voting members present are: Barbara McNeil, Carole Flamm, Jeffrey Lerner, Michael Manyak, Donna Novak, Steven Guyton. Dr. Frank Papatheofanis will vote in the event of a tie. A quorum is present. No one has been recused because of conflicts of interest.

00006

The following announcement addresses conflicts of interest issues associated with this meeting and is made part of the record to preclude even the appearance of impropriety. The conflict of interest statutes prohibit special government employees from participating in matters that could affect their or their employer's financial interests. To determine if any conflict existed, the Agency reviewed all financial interests reported by the committee participants. The Agency has determined that all members may participate in the matters before the committee today.

With respect to all other participants, we ask that in the interest of fairness that all persons making statements or presentations disclose any current or previous financial involvement with any firm whose products or services they may wish to comment on. This includes direct financial investments, consulting fees and significant institutional support.

I would now like to turn the meeting over to Dr. Sean Tunis, and Chairman Dr. Frank Papatheofanis, who will ask the committee members to introduce themselves and to disclose for the

00007

record any involvement with the topics to be presented. Dr. Tunis.

DR. TUNIS: Thanks, Janet. Just very

briefly, I wanted to thank the panelists for attending today and especially for all of the extensive preparatory work I'm sure they have all done in reading the material for this meeting, which was quite voluminous.

And other than introducing myself as the director of the coverage group and the federal liaison to this panel, I just want to continue around the table and continue introductions.

DR. PAPATHEOFANIS: I'm Frank Papatheofanis. I am on the faculty of the University of California at San Diego, and I am going to be chairing the meeting today.

DR. BURKEN: I am Mitch Burken. I am a medical officer with Sean's group in coverage and I am also an acting division director in medical and surgical services.

DR. MCNEIL: I'm Barbara McNeil from Harvard Medical School and the Brigham and Women's Hospital.

DR. LERNER: I'm Jeffrey Lerner. I am vice president for strategic planning at ECRI, and

00008

I direct our evidence based practice center, as designated by AHRQ.

DR. MANYAK: I am Michael Manyak, professional and chairman of urology at the George Washington University in Washington, D.C.

MS. NOVAK: I am Donna Novak, I am a principal with Marsh McClennan Enterprise Risk Consulting.

DR. GUYTON: I'm Steve Guyton. I'm a cardiothoracic surgeon at the Virginia Mason Medical Center at Seattle.

DR. KRUBSACK: I am Arnold Krubsack, medical director for Medicare Part B in Indiana, with Administar Federal.

DR. FLAMM: I'm Carole Flamm. I am senior consultant at the Blue Cross/Blue Shield Association Technology Evaluation Center, and I was a co-author on the technology assessment report on PET that was done as a task order

through the AHRQ evidence based practice center program.

DR. ABRAMS: Hi. I'm Jeff Abrams, I'm a medical oncologist and I work in the breast cancer area at the National Cancer Institute.

MR. KLEIN: Mike Klein, president and

00009

CEO of R2 Technology, computer aided detection for medical imaging, previously general manager for oncology for Varian Medical Systems.

DR. PAPATHEOFANIS: Great. Well, good morning and welcome to panelists and also to the audience. As you all know, this is the first time the Diagnostic Imaging panel actually will be reviewing and considering a topic in the two-year interval since we all met. You probably have tracked the Executive Committee and some of the other panels. The Executive Committee has met at least half a dozen times and has considered numerous topics. Approximately half of the panels, I think, have yet to meet or are about to meet. And we're just kicking off, so welcome.

As you know, you were chosen to serve on this panel because of various backgrounds and various levels of expertise that you bring to the table, and what hopefully Barbara, the co-chair, and I would like to see in our deliberations today is expressions of that expertise, and a lively discussion.

Obviously, it's a very contentious or potentially contentious topic that we will be reviewing. The research that has been done and

00010

the background information you have been provided is very thorough, it's technical, and it's difficult to appreciate, and so hopefully there will be opportunities for all of you to ask questions and seek clarification during this meeting.

That's all I would like to say at this point, and I'm going to turn the mike over to

Mitch Burken.

DR. BURKEN: I think the way to start the day off is to talk about what questions are going to be posed to the panel, and let's get right to it.

There is a framework, kind of a two-part framework that we're going to be using for all the questions, and the first part of this two-prong framework is to ask, is there adequate evidence to, that PET improves health outcomes under a particular situation? And then once we have answered that first question, we will go to a second question and we'll say, if so, what is the size of the effect, and there is a seven-point scale that the Executive Committee has helped lay out for us, starting from not effective; up to less effective without advantages; less effective

00011

with advantages, and those advantages might be convenience or tolerability; then going up to as effective without advantages, or with advantages; then more effective; and then a breakthrough technology.

So with that in mind, again, that being the general framework which we have used throughout several panels, we discussed PET always in the context of a comparative technology. So in this first question, we compare PET to biopsy when there is an abnormal mammogram or palpable mass, and obviously in this situation, there is presumably a high risk of malignancy, so biopsy is considered an alternative strategy.

In the second question, we take another situation where we have a lower suspicion of cancer, and we look at the difference between PET and short interval mammographic follow-up.

In the third scenario or the third question, we look to see whether PET has a role in staging as compared to axillary lymph node dissection, and once we have addressed that issue, we find another question that opens up because sentinel node biopsy has been an emerging

diagnostic technology, so we ask is sentinel

00012

biopsy versus PET, you know, versus axillary lymph node dissection, something we ought to consider.

The fourth scenario we have is looking at PET versus standard staging tests for detecting locoregional recurrence or distant mets.

And finally, the fifth question we ask is whether PET is effective or is there adequate evidence that PET improves health outcomes in determining tumor response to treatment compared to the use of conventional response criteria.

Are there any questions about the questions? Okay.

DR. PAPATHEOFANIS: Great. Thank you, Dr. Burken. We are going to follow the agenda that has been posted and I think that we will just move along to the presentation of the technology assessment by David Samson.

The other framework in addition to the one that Mitch outlined is the one that you have in your packets and I think is available at the desk in front, and that's the recommendation from the Executive Committee for evaluating effectiveness, which is an important document that the Executive Committee has been framing for the past 18 months or so, so please keep this in mind

00013

in our discussions as well. So, welcome, David.

MR. SAMSON: Thank you for inviting me. I am associate director of the Technology Evaluation Center for the Blue Cross and Blue Shield Association, and as Dr. Flamm pointed out, we are an evidence based practice center designated by AHRQ.

The assessment that we consider today can be broken down into several parts, and these are the points I will be making. I will be going over first, the review methods that we used; then I will discuss the indications, the specific ones that we considered, the first being the initial

diagnosis of breast canter; second, initial staging of axillary lymph nodes; third is detection of locoregional recurrence of distant metastasis recurrence; and the fourth being evaluating response to therapy. I will then finish up with the conclusions.

All right. Turning first to the review methods, the following topics had to do with review methods. First, I will go over what our data abstraction elements were, I will describe the study quality characteristics, I will discuss meta-analysis, the search methods, and the study

00014

selection criteria.

Here are the data abstraction elements that we looked at, first the sample size, and we also looked at the institution that the study was performed at and the dates of the study, whether the study design was prospective, retrospective or unclear, what patient selection criteria were described, the mean patient age, and the tumor size and T stage distribution, and the technique by which PET was interpreted, whether it was qualitative, quantitative, sometimes semiquantitative, and also whether attenuation correction was performed.

Some additional data abstraction elements included whether verification bias was avoided. By this we were looking for consecutive series of patients, that qualified as a yes. If there was no information about whether the patients were selected consecutively, we in most cases put a question mark to indicate that it was uncertain.

We also looked at whether the PET imaging were read blind to the reference standard evaluation, whether the reference standard was read blind with respect to the PET image. We also

00015

gave details about the reference standard test itself, whether it was histologic or had to do

with another imaging procedure with follow-up. We looked at the unit of analysis, whether it was the lesion, perhaps a region, an anatomic region for the patient. Then we gave the diagnostic performance data, the joint events of the reference standard and the test result, whether true positive, false negative, false positive or true negative.

And then the prevalence data. And throughout the presentation, when I say prevalence, that can be used interchangeably with the pretest probability of disease.

Here are the study quality characteristics that we looked at, and I'm aware that there are other, that there are a variety of sources that you can use to document study quality characteristics. The sources that we relied on were the Cochrane collaboration methods group, and a landmark paper from 1994 in the Annals of Internal Medicine by Ehrlich et al., that were guidelines for doing systematic reviews on diagnostic tests and also meta-analysis.

So one of the key things that we looked

00016

at was whether there was a valid reference standard, again, whether tests were interpreted blindly with respect to the reference standard and vice versa, whether verification bias was avoided, and verification bias having to do with whether the test results influence performance of the reference standard. We wanted a clear description of the spectrum of disease in the study sample, clear description of other patient characteristics, clear description of the test performance, interpretation and reproducibility aspects, whether the study design was prospective or introspective, and whether there was a valid design for comparing the index test with alternative tests.

These are the criteria for what we considered a higher quality study. It had to possess three qualities: First, had to be a

prospective design, had to avoid verification bias, and the study had to use blind interpretation of the PET with respect to the reference standard. These three characteristics were intended to be used for sensitivity analyses and quantitative data synthesis, and I will get into that more later.

00017

Meta-analysis was performed in this assessment. Why do meta-analysis? First of all, you can overcome small sample sizing in studies by pooling them, you can come up with point estimates for diagnostic performance, and you can systematically assess the influence of important variables that may not influence diagnostic test performance, for example, the testing techniques, patient factors and study quality.

There are several techniques in doing meta-analysis of diagnostic tests. You can perform a conventional random effects model, or a fixed effects model meta-analysis. Disadvantages of doing that is that they tend, they do not account for the dependence between sensitivity and specificity, and therefore, tend to underestimate them.

Another approach is to use the summary receiver operating characteristic curve or ROC curve. It's important when you are using a summary ROC curve approach to keep in mind whether you're doing it on a test that was interpreted qualitatively versus quantitatively; if it's a qualitative test, then you have to be careful about selecting a point on the summary ROC curve.

00018

You can produce summary ROC curves by either nonweighting or weighting by the inverse of the variance. Waiting has the advantage of giving more attention to larger studies, and again, selecting a representative point on the summary ROC curve has to be done with great caution, especially when you have a qualitatively

interpreted test.

Here are the search methods that we used. We did our electronic search of two databases, the MEDLINE PubMed, and CANCERLIT databases. Our search strategy began by looking at radionuclide imaging as a mesh term. It was exploded to get all subordinate mesh terms. And we also looked at the word positron and PET as text words, we have the intersection of those two phrases in that search strategy. And then we also kept the intersection with neoplasms.

All of these references were loaded onto a ProCite database, and the search for breast cancer. The studies that we looked at were limited to these published in English. The electronic search was conducted from January of '66 through March of 2001. We also looked at additional sources, including reference lists of

00019

key articles, current content, and expert peer reviews.

The total retrieval from this search strategy was 163 references.

Here are our study selection criteria. First we were looking for a study that was published in a peer reviewed journal as a full article, not a conference abstract. If there were multiple reports from a single institution, we limited the inclusion of studies to the largest series for the purpose of data synthesis. We wanted at least 10 patients with breast cancer, not mixed in with other types of tumors. We wanted tomographic imaging of FDG, not planar. And we had to have a correlation of the PET results with reference standard results for both diseased and non-diseased patients. There were additional indication specific criteria that we applied.

When we applied these general criteria, a total of 32 studies were included.

All right. The first indication that we reviewed had to do with initial diagnosis of

breast cancer, and there are actually to subindications, the first having to do with

00020

obviating biopsy for a suspicious mammogram or a palpable mass, and the second selecting biopsy for a patient with a low suspicious mammogram.

For all of the indications that I will be reviewing, I will first point to some clinical issues, then state the problem formulation, and then discuss the evidence review and analysis.

All right. I would like to distinguish these first two roles for PET in initial diagnosis of breast cancer. 1-A is a patient with a suspicious mammogram or palpable mass and the idea is that if PET is negative, that patient might be able to avoid undergoing a biopsy. Now, the patients who do have a suspicious mammogram or palpable mass comprise the upper segment of the biopsy population. The lower segment would be patients who were referred for biopsy for an indeterminate mammogram usually.

But the key issue is that patients who are referred for biopsy are frequently false positives in the screening process, and they end up having negative biopsies. The question in this role of using PET is whether we can improve the selection for biopsies.

The second indication here, 1-B has to

00021

do with patients who have a low suspicion mammogram, and would be referred for shorter interval follow-up. The question here is whether some of these patients might be selected for biopsy, they could have an early biopsy and early diagnosis and may benefit from early treatment. So again, the issue here is whether we can improve the selection of follow-up for biopsy.

All right. I am going to the problem formulation for indication 1-A. These are patients who have an abnormal mammogram or a palpable mass and are recommended for biopsy. The

comparison here is going to be between using a negative PET result to avoid a biopsy, versus performing biopsy on all patients.

Some of the health outcomes that are of concern, if PET is a true negative, the benefit would be to avoid the pain an anxiety of biopsy. If PET is a false negative, the harm could come from having missed or delayed diagnosis and delayed treatment.

This is a causal chain and forgive me for this small print, I crammed it together as much as I could and made it as big as I could, but I realize that you probably can't read this. The

00022

key thing though, is to recognize that there are two paths. The first path up here is using PET; the second path is not using PET. So if a patient decides to use PET to guide the decision of whether to perform the biopsy, at this point the PET would be performed, up here the PET result would be positive and the patient would undergo biopsy. In some patients the PET would be a true positive so there would be an actual tumor found. In other patients there would be a false positive and the patient would not, would have a benign mass.

If the PET is true positive, the patient would go on to getting treatment, and in the last two columns, I point out what the outcomes are in path one compared to path two, so in path two, these are all patients who undergo biopsy, and in some cases the biopsy is positive and others it's negative, so if it's positive, these patients have the benefits associated with early treatment, and the harms of pain and anxiety of biopsy in addition to any treatment side effects.

If the biopsy is negative, the benefit would be reassurance, and the harms would have to

00023

do with pain and anxiety of the biopsy.

So, the comparison between using PET if it's positive and doing biopsy in all cases, the benefits of positive PET would be the same as those in the biopsy PET. It's only when there is a negative PET would there be any difference in the types of outcomes that could occur. So if PET is truly negative, the patient could safely avoid the pain and anxiety of biopsy. If the PET is falsely negative, there would be an undetected tumor, the patient would resume the screening schedule, but may suffer from the loss of the advantage of early treatment.

All right. The specific question, as Mitch pointed out earlier is the following: Is there adequate evidence that PET can improve health outcomes when used to decide whether to perform a biopsy in patients with an abnormal mammogram or a palpable mass? And within this question we asked two subquestions. We first wanted to know if we could reach conclusions about the diagnostic performance of PET, and then we wanted to see how the diagnostic performance translates into outcomes, and whether those outcomes would be improved by using PET.

00024

So, here's the evidence we were able to find. First, I wanted to just touch on some issues dealing with the biopsy population. First of all, there is an overall prevalence of malignancy of approximately 20 to 30 percent. The upper segment as I described are patients who have an abnormal mammogram, a palpable mass, and relatively large lesions. The lower segment are patients with an indeterminate, that should be mammogram, a nonpalpable mass, or small lesions. And for this lower segment of the population, we don't have any diagnostic performance data for PET. It's only for the upper segment for which we have any PET diagnostic performance data.

There were a total of 13 studies with a pool of 606 patients. Unit of analysis in three studies was lesion, for 191 patients. The unit

was patient for 10 studies, and 415 patients. There were consistent study selection criteria, as I described in the problem formulation, and the average tumor size across these studies was between 2 and 4 centimeters, so these are fairly large tumors.

Here is a summary of study quality characteristics. 9 of the 13 studies were

00025

prospectively designed. 3 out the 13 avoided verification bias. 7 clearly indicated that PET was read blind to the reference standard, and none of the studies indicated whether the reference standard was read blind to the PET.

Here is a summary of the diagnostic performance data. In individual studies, the range of sensitivities was between 79 and 100 percent. The random effects meta-analysis comes up with a point estimate of 88 percent and a confidence interval here between 83 and 92 percent. Specify ranged between 50 and 100 percent, with a random effects meta-analysis point estimate of 79 percent and a 95 percent confidence interval between 71 and 85 percent.

Here is the graphic of the meta-analysis, and each line here represents an individual study, and the random effects meta-analysis point estimates are down here at the bottom.

Here is the summary ROC curve and as I said earlier, you have to be careful in using a random effects meta-analysis because it tends to underestimate the diagnostic performance, because it doesn't account for the dependence between

00026

sensitivity and specificity, and this can be seen in any summary ROC curve to the extent that the X here which represents the random effects meta-analysis point is below the summary ROC curve. And the curve that we used was the one that was weighted by the inverse of study

variance.

So, the random effects meta-analysis doesn't underestimate the sensitivity and specificity by a great deal, it's pretty close to the curve. But we decided that just to eliminate the underestimation of diagnostic performance with a random effects meta-analysis, we chose the point on the summary ROC curve nearest to the random effects meta-analysis point. And we did that partially because we wanted, you could ideally select any point on the summary ROC curve and that would represent the diagnostic performance of PET. However, we think that the advantage of doing a point near the random effects meta-analysis point is that it represents an average diagnostic performance.

And you could say that you would be looking for points on the curve that have higher sensitivity. However, you could only do that if

00027

you could realistically adjust your criteria for a positive test result, and when you're doing a qualitative test, that's very difficult. So we decided to look at this point here on the curve closest to the random effects meta-analysis point as being a good representative choice.

We did plan to do sensitivity analysis, but only one study met study selection, or the quality criteria, and so we didn't go through with that.

The analysis of outcomes can be done from two different perspectives, and I will be walking you through some examples to try to make this clear. The first perspective is that of the population, so using a given prevalence and estimates of sensitivity and specificity, as well as the causal change that I talked about earlier, we can calculate the probabilities of outcomes before the PET scan results are known.

Now, from the perspective of a patient who has a negative PET scan, the perspective is different, but using different given prevalence,

in other words pretest probability, and the same information here, we want to calculate the negative predictive value or the post-test

00028

probability, and the associated probabilities of outcomes for a patient with a known negative PET scan.

Now, from the population perspective the question that you would ask a patient would be this. Based upon the probabilities to follow, would you be billing to let the results of PET guide your decision to undergo biopsy? That is, if PET is positive, do the biopsy, if it's negative, skip biopsy. The alternative to using PET to guide the decision is for all patients to undergo biopsy.

Now we know the probabilities before you undergo the PET scan and that's all based on the diagnostic performance estimates and prevalence. Now, the two examples that I will be using will be first with a prevalence of 50 percent and second with a prevalence of 75 percent.

Now for a, the perspective of a patient who has a negative PET scan, the question is this: Based on the probability of PET missing a cancer, would you still be willing to skip the biopsy if your PET scan is negative. The probabilities of true negative and false negative differ in this

00029

perspective from that of a population, because the denominator is different.

Now, although there is, this is described as a known negative PET, we know the probabilities before you undergo the PET scan, and you can imagine making the decision, so we don't actually have to put the patient through the PET scan and come out with a negative result in order to go through this scenario. And again, the two examples I will be using are prevalence of 50 percent and 75 percent.

All right. This is the first example. Prevalence is 50 percent, here's the two-by-two table, we are assuming there is a total population of a thousand individuals. This column represents patients who have malignant lesions, these patients have benign lesions. This row is patients who test positive on PET and this row is for PET negative patients.

So here is the sensitivity and the specificity, 89 percent and 80 percent. This is the point on the summary ROC curve closest to the random effects meta-analysis point. And here are the probabilities of the different events. We have the true positive, false negative, false

00030

positive or true negative. So when the prevalence is 50 percent, the probability of a true positive result is 44.5 percent, the false negative probability is 5.5 percent, the true negative is 40 percent, and the false positive is 10 percent.

Now, you will see that in this column, I do it from the population perspective and in this column I do it from the PET negative individual perspective. And so, the two outcomes that we're going to be most interested in are the false negative and true negative, and from the population perspective, these are what the probabilities are. However, when you get to the perspective of a patient testing negative on PET, the probabilities for false negatives and true negatives change, and the reason is that you have a different denominator. The denominator from the population perspective is the total of all the cells of the two-by-two table, whereas from the perspective of an individual with a negative PET scan, the denominator is only the row marginal total for the PET negative patients.

So, the risk of false negative rises as you go from the population perspective to the individual perspective.

00031

On this slide I summarize what is already in tables 3 and 4 of the document, and at a prevalence of 50 percent, these are the probabilities. Now I, the first two columns represent the population perspective and the third is the individual perspective. So the outcomes, if PET is true positive or if the patient is being managed in the path in this which all patients would undergo biopsy, that would be a positive biopsy. The benefit would be whatever outcomes would be associated with the appropriate treatment, and the probability of having this outcome would be 50 percent.

If the biopsy was the choice, and the harm of having either a PET false positive or a negative biopsy would be the morbidity associated with biopsy, and that would also be in 50 percent.

The two key outcomes that we're interested in are the two in the center here, the harm associated with the false negative PET, which could possibly result in late treatment, or the benefit of a true negative PET, in which the patient could avoid the morbidity of biopsy. So the patient could look at these numbers and decide whether the benefit that you gain in terms of the

00032

probability of avoiding the biopsy morbidity is worth the harm that you get from delaying treatments. And so, the risk-benefit trade off would take into account these results, first from the population perspective.

Once the patient has a negative PET result, the probabilities change, so the risk of a false negative, having delayed treatment would rise to 12.1 percent, and the benefit would be about 88 percent.

Now, this is the second example on which the prevalence is 75 percent, the sensitivity and specificity are the same as in the previous example, 89 percent and 80 percent. The probabilities of a true positive are 66.8 percent, false negative 8.2 percent, true negative 20

percent, and false positive 5 percent.
From the perspective of a patient who had a negative PET scan, the probabilities differ again, because the denominators differ, so the false negative risk goes from 8.2 percent at the population perspective to 29.2 percent at the individual perspective, and I think most people would agree that the risk-benefit trade-off is not an acceptable one with these kind of numbers.

00033

Again, I present the same information here, this can be found in tables 3 and 4 of the document. It's the same as on the previous slide, just presented with descriptions of what the outcomes are. So again, we're comparing the harm of delaying treatment with the benefit of avoiding the morbidity of biopsy, and you have to balance the 20 percent benefit with the 8.2 percent harm from the population perspective, and versus the individual perspective of a 70.8 percent benefit against the 29.2 percent.
All right. Our conclusions are that the diagnostic performance data that are available apply only to the upper segment of the biopsy population, not to the lower segment, so there is incomplete data for the full spectrum of patients that we might be interested in.
Only one study met all of the criteria for a higher quality study; the sensitivity estimate was 89 percent, specificity was 80 percent. For the intermediate to higher prevalence spectrum, the risk-benefit trade-offs do not appear to be acceptable.
All right. Turning to the indication 1-B, having to do with initial diagnosis of breast

00034

cancer, the problem formulation is this. The patients of interest are those who have low suspicious findings on mammography and other routine imaging procedures that are referred for short interval follow-up, from three to six months

in frequency. The comparison we're using here is using PET to elect early biopsy or avoid short interval follow-up, versus doing short interval follow-up in all patients.

The health outcomes associated with different PET results, if PET is true positive, it could lead to earlier detection and treatment of malignancy. If PET is true negative, patients could forego short interval follow-up and revert to a normal screening schedule, so they would be avoiding some inconvenience. The false negative PET outcome would entail foregoing short interval follow-up and the potential benefit of earlier detection and treatment. And the outcome associated with the false positive PET would be the morbidity associated with biopsy.

The specific question that we're asking here is, is there adequate evidence that PET can improve health outcomes by leading to earlier and more accurate diagnosis of breast cancer, compared

00035

to short interval mammographic follow-up, in patients with a low suspicious finding on mammography or other routine imaging procedures. And again, within this question, we're asking whether we can reach conclusions about diagnostic performance of PET and can the use of PET improve the outcomes by selecting follow-up or biopsy.

What is the evidence? Well, there are no studies available, so we can quite quickly reach the conclusion that we don't know what the diagnostic performance data or health outcomes are.

All right. Turning now to the second indication, this is the initial staging of axillary lymph nodes, again, we going to be looking at clinical issues, the problem formulation and the evidence review.

The clinical issues, the patients who are undergoing staging of axillary lymph nodes by PET or some other noninvasive procedure are undergoing that testing in order to determine

whether they might need to undergo axillary lymph node dissection. And the roles of axillary lymph node dissection could be either to define prognosis, to guide treatment decisions, and it's

00036

also wondered whether the procedure itself is therapeutic. It might contribute to local control of the tumor as well as, there is some question about whether it improves survival, although the data has not demonstrated that yet.

But the key thing that we are focusing on here is guiding treatment decisions and in particular, a patient who has a positive lymph node on pathologic analysis, an axillary lymph node dissection, would be a good candidate to undergo adjuvant therapy. Now this is complicated by the fact that some patients who are negative on axillary dissection would also, may choose adjuvant therapy.

Here are some of the outcomes that we have been able to identify that are associated with adjuvant therapy in patients who are either lymph node positive of lymph node negative. So, patients will undergo either adjuvant chemotherapy or hormonal therapy. The median overall survival increases by two years, and ten-year overall survival, there is a difference between patients who get adjuvant therapy and those who don't at ten years, of 6.8 percent.

In patients who are lymph node

00037

negative, the chemotherapy can have a significant advantage for ten-year overall survival but it's a smaller one, it's 3.5 percent. Patient preferences can play a big role in whether a patient chooses adjuvant therapy, and different patients may value the survival benefits of adjuvant therapy in different ways, and other patient may value the adverse effects of adjuvant therapy and so may make different decisions.

Sentinel node biopsy is an emerging

technique that is used for a similar purpose compared to PET for staging axillary lymph nodes. It s an invasive procedure, however. The technique involves using either a blue dye or a radiotracer injected near the tumor site, and either the dye or the tracer is tracked to determine which is the first lymph node that is visualized or localized. That would be called the sentinel node. And if it's positive, that patient may go on to full axillary lymph node dissection. If it's negative, patients might be able to avoid the full axillary dissection.

As the issues in evaluating sentinel node biopsy, we're looking first of all at sensitivity. So a false negative sentinel node

00038

would be one in which the node would be negative, but other downstream nodes might be positive, and that would be considered a skipped metastasis. The specificity for sentinel node biopsy is always 100 percent. Each positive result from a sentinel node biopsy is pathologic positive, so it's, there is no possibility of a false positive.

We did a systematic review of 21 studies in over 3,000 patients and the results we got were a weighted average rate of successful localization of 90.1 percent and a random effects meta-analysis point estimate for sensitivity of 89 percent, and the confidence interval was between 86 and 91 percent.

The problem formulation that we used in this indication, the patients that we were concerned with are those patients who have confirmed primary breast cancer, no palpable axillary lymph nodes, and no evidence of distant metastasis. The comparison we're using here is between using PET to decide whether to perform axillary lymph node dissection versus performing axillary lymph node dissection in all patients.

The key health outcomes of interest are when PET is a true negative, the patient could

00039

avoid the complications of axillary lymph node dissection; when PET is a false negative, that patient, if the result is used to avoid adjuvant chemotherapy or other treatment, that patient would have an undetected positive lymph node and could be considered undertreated.

And again, the causal chain is in very tiny print and I will try to walk you through it. Again, we have two paths. The first is using PET to select whether to undergo axillary lymph node dissection, and the path down here is using, is not using PET, so all patients would undergo axillary lymph dissection.

And again, the outcomes associated with doing axillary node dissection in all patients are down here, and the outcomes associated with using PET to choose axillary lymph node dissection are up here, and are viewed in comparison with this path. So we are interested in up here the kinds of outcomes that differ in this path from this path, and I'll get into that in a moment.

But anyway, if the axillary lymph node dissection reveals positive lymph nodes, the causal chain here assumes that patients would the initiate adjuvant therapy and the outcomes would

00040

be those associated with adjuvant therapy. If there are no positive lymph nodes found, then the patient would not elect adjuvant therapy and would just undergo monitoring for recurrence. And the outcome, the benefit of the negative PET scan -- I'm sorry, negative axillary node dissection -- would be the prognostic information that it supplies. And if the axillary node section is positive, the harm would be the adverse effects associated with axillary lymph node dissection and with adjuvant therapy. And for those patients who are lymph node negative, the harms would be the adverse effects of axillary node dissection.

So, if the patient decides to use PET to guide the choice in whether to have axillary

node dissection, if it's positive they would undergo node dissection, either the PET was truly positive or false positive. If it's truly positive, they would be getting adjuvant therapy and the benefits would be the same as here on this path. If the PET is falsely positive, the patient would have the adverse effects of axillary dissection. If PET is negative and skips axillary dissection and it's truly negative, they would benefit by avoiding the adverse effects of

00041

axillary node dissection. If PET is falsely negative, then they wouldn't be getting adjuvant therapy and they would be undertreated.

The specific questions that we asked are, is there adequate evidence that PET can improve health outcomes when used to decide whether to perform axillary lymph node dissection. And again, we wanted to know whether we could get conclusions about the diagnostic performance of PET and whether use of PET to decide whether to perform axillary node dissection could improve outcomes.

And a second question is whether there is adequate evidence on the previous question, should we do be doing a more detailed analysis of sentinel node biopsy versus PET, as alternatives to actual lymph node dissection.

Here's the evidence that we were able to find. First, I want to go over some issues dealing with population. You can break down patients who undergo PET into those who have palpable axillary lymph nodes versus nonpalpable axillary lymph nodes, and the disease spectrum in those groups, if they're palpable, these are patients who have larger metastatic foco in lymph

00042

nodes, and patients with nonpalpable nodes would have smaller foci.

There are potential differences in the diagnostic performance of PET for these two

segments letters of the population, and axillary lymph node dissection would probably be likely for patients who have palpable axillary lymph nodes regardless of imaging. So, we are really interested in the patients who have nonpalpable axillary lymph nodes, because those are the patients for whom use of PET really could make a difference in determining whether they have axillary lymph node dissection, and it's fundamental to assess the diagnostic performance of PET for the patients who have nonpalpable axillary lymph nodes.

All right. We came up with a total of four studies and 269 patients who had nonpalpable axillary lymph nodes and there was specific data on the diagnostic performance of PET for those patient. In the appendix of the document we actually list a larger group of studies in which the evidence is presented irrespective of whether the patients had palpable or nonpalpable lymph nodes.

00043

Here are the study quality characteristics. Four of the four studies were prospective designs. One of them avoided verification bias. Three out of four read PET blind to the reference standard, and none of the four read the reference standard blind to PET.

Here is the summary of the diagnostic performance data. In the four studies, sensitivity ranged between 40 percent and 93 percent. The random effects meta-analysis comes up with a point estimate of 80 percent, and a 95 percent confidence interval of 46 to 95 percent. That's really quite large.

The specificity ranged between 87 percent and 100 percent. The random effects meta-analysis point estimate was 89 percent, with a more narrow confidence interval between 83 and 94 percent.

Here is the graphic representation of the random effects meta-analysis, with the point

estimates of sensitivity and specificities at the bottom of the graph.

Here is the summary ROC curve. Now I should throw in a note of caution that doing a meta-analysis on such a small body of evidence is

00044

an exercise that you might question and with good reason. I think we went through this exercise just for illustrative purposes. I think the key point was that there was a very large confidence interval around the sensitivity and ultimately, we would conclude that there is not sufficient evidence to estimate diagnostic performance for such a small group of studies.

But, if you go through the exercise, this is what the summary ROC curve looks like. When the curve is weighted by the inversive study variance, it's the one on the inside here. The X represents the random effects meta-analysis curve. If you choose the point nearest on the summary ROC curve, the sensitivity and specificity estimates are here, so the sensitivity would be 81 percent, specificity would be 95 percent.

A sensitivity analysis was not possible with respect to study quality.

All right. Again, we're looking at the outcomes from two perspectives, first the population perspective, and the question we would ask the patient would be, based on the following probabilities, would you be willing to let the results of PET guide your decision to undergo

00045

axillary lymph node dissection? If PET is positive, do the axillary lymph node dissection; if it's negative, skip the dissection. The alternative to PET guiding the decision is for all patients to undergo axillary lymph node dissection. We know the probabilities before the patient undergoes the PET scan, and the two examples that we're going to be using are a prevalence of 30 percent and a prevalence of 50

percent.

From the perspective of a patient who has a negative PET scan, the question is, based on the probability of PET missing a positive axillary lymph node, would you still be willing to skip axillary lymph node dissection if you had a negative PET scan? The probabilities of true negative and false negative differ from the population perspective because the denominators differ, and we know the probabilities of the PET scan before we actually undergo the procedure. And again, the two examples are prevalence of 30 percent and 50 percent.

The two-by-two table is similar to the first ones I presented on detection of breast cancer, with the exception that the columns

00046

represent whether axillary lymph node dissection as the reference standard, comparing positive lymph nodes versus negative lymph nodes and for this case, these are again using 100 or 1,000 patients as the example, at 30 percent prevalence, these are what the cell counts would be. The sensitivity, again, would be 81 percent and specificity would be 95 percent. The probability of a true positive result would be 24.2 percent, false negative result would be 5.7 percent, true negative result would be 66.5 percent, and a false positive would be 3.5 percent.

Now, as you go from the population perspective to the perspective of a patient with a negative PET scan, the probabilities of false negatives and true negatives change because the denominators change. So, at the population perspective, the denominator is 1,000; at the PET negative perspective, this is the denominator. And so, the false negative risk goes from 5.7 percent to 7.9 percent.

And here, I present the evidence in the same form that's shown in tables 9 and 10 of these documents. At a prevalence of 30 percent, these are what the probabilities are. Here are the

00047

outcomes. If PET is true positive or if the patient chooses to go straight to axillary lymph node dissection and that's positive, the outcomes would be associated with choosing adjuvant therapy, and since the prevalence is 30 percent, the probability would be 30 percent of that outcome. For false positives on PET or having a negative axillary lymph node dissection, the outcomes would have to do with the morbidity of axillary node dissection and the probability would be 70 percent.

The key outcomes to look at are in the center here. If PET is falsely negative, the outcome would be the loss of the benefit of adjuvant therapy so it would be undertreatment. If PET is truly negatively, the patient would safely be able to avoid axillary lymph node dissection and its morbidity.

So, we're trying to decide whether the benefit outweighs the harm. The risk of undertreatment is 5.7 percent from the population perspective, compared to a benefit of 66.5 percent of avoiding the morbidity of axillary lymph node dissection, but when you go to the individual perspective, the risk of fall negative rises to

00048

7.9 percent, and in this case and in the next case, we conclude that that trade-off is not going to be judged as acceptable to patients.

Here is the second example where the prevalence is 50 percent, again, sensitivity is 81 percent, specificity is 95 percent. These are the calculations for the probabilities of the different outcomes from the population perspective and the perspective of an individual with a negative PET scan.

And here again, we present the information as it is in tables 9 and 10, and the key thing to look at is whether the trade-off between the benefit of avoiding axillary lymph

node dissection morbidity and undertreating is an acceptable one. And a risk at the population perspective of 9.5 percent is pretty high and would probably be unacceptable to patients. But when you go to the perspective of an individual with a negative PET scan, the false negative risk is 16.7 percent, which is quite high.

All right. The conclusions that we reached here, first of all, the diagnostic performance data applicable to the nonpalpable population is sparse. There were four studies and

00049

269 patients. Sensitivity was 81 percent, specificity was 95 percent. Even if you could have greater confidence in the diagnostic performance data, in the intermediate prevalence spectrum the risk-benefit trade-offs do not appear to be acceptable.

All right. Let's move on to the third indication, and this is detection of locoregional recurrent or distant metastasis recurrence. I will look at the background issues, the problem formulation and the evidence review. The clinical issues here have to do with whether the patient is undergoing local versus systematic therapy, PET might influence the choice of that. There might be more accurate information from PET which could lead to early detection of recurrent metastasis. There might be improved timing or improved choice of treatment.

The kinds of studies that we're looking at that we want to see are comparative studies, so these are studies in which PET and some other kind of imaging test is performed on the same group of patients, and both of those tests are compared against a reference standard. We want to have information on the discordance and concordance

00050

between PET and alternative tests. We want to know the frequency with which each test is correct, when discordant, and the frequency with

which one test or the correctly upstages or downstages the disease when it's added to other tests. The key thing here is that it is crucial to have comparative studies.

The reference standard in studies in which you're looking for metastasis or recurrence is not as clear-cut as it is when you're doing an initial workup. So when you're doing an initial staging of lymph nodes or your initial detection of the primary tumor, you almost always can get a histologic reference standard. However, when you're doing imaging for recurrence or distant metastasis, it's usually not feasible to biopsy widely, so in many cases, you would have instead of a pathologic reference, you would have some kind of follow-up study, and the key thing here is to have an adequate duration of follow-up.

The bottom line is that there should be a more flexible approach to what you would accept as a valid reference standard for studies in which you are looking at recurrence or metastasis.

Here is the problem formulation. The

00051

patients are patients who either have locoregional, might have locoregional recurrence, and these might be symptoms referable to the brachial plexus, or patients who are suspected to have distant metastasis, and this could be either in the setting of initial staging or after treatment. The comparison is between PET and routine tests, including physical examination, chest x-rays, CT, MRI, radionuclide bone scanning, and we would be making comparisons by anatomic site.

There are two comparisons that could be performed. First, PET as an adjunct to other tests so you're adding PET to other tests, or PET done as a replacement for other tests. The health outcomes that we're interested in, if PET is correct, the patient could receive initial follow-up treatment appropriate for that stage, they might receive earlier initiation of treatment

and avoid the morbidity of unneeded treatment. If PET is incorrect, patients may undergo unneeded biopsy and potential harmful and unnecessary treatment and may forego the potential benefits of timely initiation of treatment.

Here's the specific question. Is there

00052

adequate evidence that PET improves health outcomes as either an adjunct or a replacement to the standard tests in detecting either locoregional occurrence or distant metastasis recurrence. We want to know the conclusions about the diagnostic performance and also whether use of PET in altering patient management improves health outcomes.

Here's the evidence. With respect to locoregional occurrence we have two studies, and the evidence is really quite meager. There is ten patients from Hathaway; these are patients who were referred because of signs or symptoms occurring in the axilla or nearby. The sensitivity for PET was 100 percent, for MRI it was 56 percent, but with a study this small you can't put a lot of confidence in these numbers.

The Bender study included 75 patients and they selected patients based on having suspected recurrence or systemic disease in patients who are equivocal on other imaging tests. Now while this was a larger study, I think the key thing here is that there was a major concern about the reference standard that they used. The authors claimed that they did a histologic

00053

reference standard in I think 90 percent of patients, but they presented data for not only locoregional sites but a number of other sites, and it's really quite unlikely that they did histologic sampling for large numbers of patients who had no recurrence or metastasis.

So, I don't think you can put any faith in these estimates of sensitivity and specificity.

However, at local site and at lymph nodes, PET had lower sensitivity at the local sites compared to CT or MR, and comparable specificity. When you looked at lymph nodes, PET was more sensitive than CT or MR, with comparable specificity.

But overall, this study has had a major problem with what the residence standard was and it calls into question some of these findings.

Looking at distance sites, there were five studies with a total of 196 patients. We did a site specific analysis and the most evidence that we had was on bone. Here are the study quality characteristics. First, three out of the five were prospective. None of them avoided verification bias. Three read PET blind to the reference standard, and none read the reference standard blind to PET.

00054

With respect to detecting bony metastasis, the Lonneux study included 11 patients, really small sample. There were no false negatives and one false positive. The Bender study, the same one I just discussed, the major problem with the reference standard in this case, so I am not even going to discuss the diagnostic performance right now.

Probably the best study is the Schirrmeister study, 34 patients. PET had a sensitivity of 100 percent, compared to bone scan 83 percent, and it was also more specific, 94 percent for PET and bone scan had a 69 percent specificity.

The Cook study included 23 patients and only reported data on the mean number of lesions detected. PET detected more lesions per patient than bone scan.

The Mortimer study looked at whether PET could detect bone metastasis earlier than other imaging techniques, and it did so in two patients.

So overall, this body of evidence is insufficient to reach conclusions about diagnostic

performance.

00055

There were three studies that gave us data on liver metastases. The Lonneux study discussed six cases, there were five true positives and one false positive. The Bender study, again, had the reference standard problem. There were only two liver metastases in the whole study. And in the Mortimer study, there was one liver metastasis, so that evidence is inadequate.

On lung metastases, a similar situation. The Lonneux study reported four true positives and one false positive, and in the Bender study, there were six lung metastases.

So the conclusions overall for indication number three are that the data are sparse, five studies all together, 196 patients, there are no data available on results that are either discordant or concordant, and no data on the frequency of which test is correct, when the results are discordant, and no data on the frequency of correct upstaging or downstaging.

I would throw in a caveat that we did get a very recent study published by Huebner this month, I didn't get it until last Friday, and I have some information about it but I don't think it adds anything. It does actually give a little

00056

information about discordance and concordance but as I said, it doesn't change the conclusions.

The final indication that we addressed is PET for evaluating response to treatment. The problem formulation is this. Patients are those undergoing multicourse treatments. The comparison is between PET and routine tests, which can vary by treatment type but can include physical examination, mammography, x-ray, CT, MRI, and bone scan.

These are health outcomes. If PET is correct, you might be able to initiate new treatment, continue effective treatment,

discontinue ineffective treatment, and identify disease free patients for continued monitoring. PET might improve the timing of treatment decisions by either allowing earlier discontinuation of ineffective treatment or earlier initiation of a new treatment, and if PET is incorrect, the consequences include continued harmful side effects that might affect the treatment, or foregoing the benefits of additional treatment.

This is the question we addressed. Is there adequate evidence that PET can improve

00057

health outcomes by providing either a more accurate or an earlier determination of tumor reasons to treatment compared with the use of conventional response criteria, which may rely upon clinical exam or other imaging tests. We wanted to know about diagnostic performance and outcomes.

There were four studies all together, for a total of 103 patient, and they looked at different treatment regimens. Mortimer used hormonal therapy, that was tamoxifen. Schelling and Smith both used chemotherapy but different measurements. And the Wahl study was a combination therapy, chemo and hormonal therapy.

Here is some study quality characteristics. All of them were prospective designs. None of them avoided verification bias. One out of four read PET blind to the reference standard, and none of them read the reference standard bind to PET.

All right. This is a busy slide, but I'll try to walk you through it. The Mortimer study included 40 patients who had hormonal therapy, tamoxifen. The PET result they were looking at was a specific change in PET at seven

00058

to ten days after treatment, and they got a sensitivity and specificity of 95 and 89 percent.

The Schelling study selected 22 patients who got epirubicin and cyclophosphamide or epirubicin and paclitaxel. They looked at the PET results correlated with the conventional response criteria either at the end of the first course for 16 patients or at the end of the second course for 22 patients, and the results differed depending on when you did the PET scan, and also for how many patients were included. So it raises the issue of just how much faith you can put into a specificity of 100 percent when not all patients were included.

The Smith study included 30 patients who had chemotherapy; this was cyclophosphamide, vincristine, doxorubicin, and I can't remember what the P stands for. Or they had docetaxel. The results of the PET scan were correlated with the results of pathologic findings at the time of surgery, so these are patients who were undergoing actually a neoadjuvant chemotherapy and the criteria for response differed in these two cases. In the first case we were looking at patients who either had a pathologic partial response or a

00059

pathologic complete response, and in the second case only patients that had a pathologic complete response. And the PET result they were looking for was at least a 10 percent decrease in the quantitative PET index in the first case or at least a 20 percent decrease in the quantitative PET index in the second case. So, depending on what your definition of the reference standard response is and the definition of the PET response is, you get different estimates of diagnostic performance.

The final study was by Wahl, 11 patients. These were patients who had cyclophosphamide, doxorubicin, methotrexate, fluorouracil, tamoxifen, and Premarin. This is a nonstandard treatment regimen. The PET was looked at after the first course of treatment, and it was 100 percent sensitivity and specificity.

Overall, this is a small body of studies, each of them had small numbers of patients, and it's a fairly heterogenous group of studies, different treatment regimens, and they evaluated the evidence in very different ways. So the conclusions are that the studies are heterogeneous, the data is sparse and

00060

insufficient, and the potential for undertreatment is substantial. So wherever there is a false negative PET, these are patients who could possibly be withdrawn from effective treatment.

The overall conclusions for the technology assessment are as follows.

For indication number one, we have diagnostic performance data applicable to the upper segment of the biopsy population but not to the lower segment, so we have incomplete data on a full spectrum of patients. One study met study quality criteria; sensitivity was 89 percent, specificity was 80 percent. For the spectrum of patients who had intermediate to higher prevalence, the risk-benefit trade-offs do not appear to be acceptable.

In indication number two, diagnostic performance data that is applicable to nonpalpable axillary lymph node population is sparse. Poor studies, 269 patients. Sensitivity estimate is 81 percent with a wide 95 percent confidence interval, specificity is 95 percent. Even if we had greater confidence in the diagnostic performance data, in this intermediate spectrum of prevalence for positive lymph nodes, the

00061

risk-benefit trade-offs do not appear to be acceptable.

For indication number three, the data are sparse and insufficient, five studies, 196 patients. No data until just recently about concordance or discordance or the frequency with which PET is correct, when it's discordant with

other types, and the frequency with which PET can correctly up or downstage patients.

Patient indication number four is represented by heterogeneous studies, they are few in number and have a small pool of patient sample. And again, the potential for undertreatment is substantial.

Thank you for your attention. At this point, I'm done.

DR. PAPATHEOFANIS: Thank you, David, that was a great presentation. For the panel, this is our opportunity to ask questions of David. Will you be here all day or what is your plan?

MR. SAMSON: I'm here all day.

DR. PAPATHEOFANIS: So if you don't get your chance now, we will bring him up again later.

DR. KRUBSACK: I have a couple questions. I'm trying to put the assessment in

00062

context, and let me ask a couple questions in that regard. What is the harm of a negative biopsy on the ability of future mammograms to detect and exclude disease, that is to say how much disease will be missed in the future or how many more unnecessary biopsies are going to result from scar tissue resulting from previous biopsy?

MR. SAMSON: I don't have a good answer to that. I think that would be extremely difficult to quantify. I know that that has been raised as an issue, that performing a biopsy changes the architecture of the tissue and can make it difficult to find new disease, but to quantify the risk associated with that would be extremely difficult.

DR. KRUBSACK: Second question. This is in regard to PET false positives. Could the PET false positive actually be a true positive? That is to say, what studies exist on biopsy to show that a negative biopsy might actually be a false negative biopsy, or said in a different way, what studies demonstrate that breast biopsy is 100 percent reliable, or said in a different way, what

studies do we have that a biopsy always acquires the tissue in question?

00063

MR. SAMSON: I don't think that there is any perfect reference standard when you're evaluating diagnostic tests. I know that there are problems with the sensitivity of the reference standard itself. But when you're comparing PET with other tests, you have to choose a single reference standard to judge all tests by. And in the case of indication number one where we're looking at using PET versus performing biopsy, what would the alternative be to doing biopsy? Would it be mastectomy? I don't think we have a good answer to that question.

DR. KRUBSACK: Yeah, but my question really is, how gold is the gold standard?

MR. SAMSON: It's as gold as it can be. I don't think there is any alternative to biopsy other than mastectomy, and that's not realistic.

DR. KRUBSACK: Okay. The last question is in regard to the technology of the PET. You know, PET is an emerging technology and the expectation might be that the studies that we're looking at could vary in the state of technology that is used. And this would significantly impact the sensitivity and specificity. So when you did the assessment, did you make any effort to

00064

evaluate the studies on this basis, and then weigh these results appropriately?

MR. SAMSON: The one variable that we did look at was whether the studies used an attenuation correction. We didn't look at anything more specifically than that. We didn't do a formal sensitivity analysis by attenuation correction, but the -- and I don't have the number at the top of my head on how many of the studies did attenuation correction, but it was the vast majority, so it's unlikely that it would have been informative to do a sensitivity analysis by

whether attenuation correction was done. I think when we eye balled it, the results didn't really show any pattern of better or worse results.

DR. PAPATHEOFANIS: Michael?

DR. MANYAK: I actually had the same thought about the gold standard issue that was brought up by my colleague over here, and I would think that would also carry out certain lymph node dissections where you are using sentinel node biopsy. In other areas of cancer, we know that there are skipped lesions, and I don't know the incidence with breast cancer but that is something inherent to this kind of comparison, there is a

00065

problem with that as a gold standard, so again, we may not have the data. But to hold, to use that as the absolute comparator is something I don't know how to get a handle around. You have already answered the question, so I am just raising this again.

I had one other question also, and that is maybe for medical oncology colleagues on the panel. What is truly the effectiveness of adjuvant therapy for breast cancer for positive axillary lymph nodes? And the reason I ask it is because, what is the consequence of the false negative test in reality for the patient? And I need some guidance on that because I am not a medical oncologist, breast cancer is not my particular field. So maybe someone could shed some light on that.

DR. ABRAMS: I will take that one. I think the statistics quoted that are largely based on the meta-analysis that came out of the groups that Oxford has performed, a meta-analysis on all the adjuvant trials done worldwide in breast cancer, would show that for node-positive disease at about 10 years, and the results now go out to even 15 years, there is about an 8 percent or so

00066

improvement survival, and depending on how you

want to, if you want to take the medians of the curves of the survival curves, you can show that to be on average maybe a two-year difference in median survival in patients who take adjuvant therapy with positive nodes versus those who do not receive it. So I think that's where those numbers came from, and I think that's probably the best data that exists right now on that question.

DR. MANYAK: And the mortality rate from the chemotherapy regimens this day is?

DR. ABRAMS: It's under 1 percent. There is some mortality, especially -- it goes up a little bit with adriamycin containing regimens because there is some slight degree of heart failure and there are some low instances of leukemias induced by chemotherapy. Those would be the major treatment induced causes of late term mortality.

There can be infections short term and also rare, way under 1 percent, so there is some trade-off, but it's, you know, versus the 8 percent gain and under 1 percent mortality, still comes out on the benefit side, and actually, that 8 percent, you can calculate it off the negative

00067

effects, so it was taking that into account.

DR. MANYAK: Thank you.

DR. PAPATHEOFANIS: Barbara?

DR. MCNEIL: Can I ask Jeff a question, going back to the false negative issue regarding axillary nodes and the gold standard. It would strike me that the false negative issue for the reference standard or the lack of 100 sensitivity in the reference standard really doesn't apply very much in that particular situation, because as I would understand it, if the patient were having axillary node dissection or a set of sampling, there would be several samples, so the chance that all of them would be falsely negative, you just keep multiplying out and the probability gets to be vanishingly small. So I think when we think about tarnished gold standards, which we perhaps

want to do, it probably does not apply to the axillary node area because we are just multiplying out, we're increasing the chance every time we do another section within a node, or we section more nodes, that we're going to get a hit. Is that true?

DR. ABRAMS: I think it's true what you say that when you have 20 nodes to look at and the

00068

pathologist takes one section of each, they increase their chances of finding it. But what the sentinel node procedure has taught us is that if you study one node very closely and do 20 slides through that note node and then use special techniques, you can sometimes find things that you didn't find in the 20. So it cuts both ways notice sense that you know, how -- you have to realize there is a sampling error in pathology and even with small biopsy samples, they take a few slides and they feel statistically they have a pretty good chance of finding something if it's there, but you'd have to take many many more slices if you wanted to get that risk down, if you wanted to make that gold standard as pure as it could be, and that is weighed against the ability to get all the work done that we have to do.

So, there have been studies that have looked at ding 20 slices in every node, and they do find a little bit more, so there is a false negative rating in pathology.

DR. MCNEIL: Could I just follow up on the gold standard because in some ways I'd like to get it off the table. It seems to me we have to live with what is our tradition of medicine and we

00069

have to go by a gold standard, and maybe it needs a little polish, but it's probably as good as we can do.

Are there any examples, I guess Jeff, you're the oncologist to answer this, in medicine where patients would be treated definitively for

cancer on the basis of a positive say screening test, which is what we're talking about here, and a negative biopsy? Does that ever happen?

DR. ABRAMS: I hate to say never, but I can't think, especially when you use the word screening, as opposed to the more metastatic disease and all that, in screening I would have to say no, I think people there, the standard is to have a positive biopsy at this point.

DR. MCNEIL: So would it be reasonable then to get the gold standard issue off the table for discussion of these issues?

DR. PAPATHEOFANIS: I think we're done with the gold standard. Go ahead, Jeff.

DR. LERNER: Just one thought on that. Anything we want to say about needle biopsies?

DR. PAPATHEOFANIS: You mean likelihood of sampling error in different types of biopsies?

DR. LERNER: Exactly.

00070

MR. SAMSON: In the studies that I reviewed, the needle biopsy was not performed as a reference standard in these studies.

DR. PAPATHEOFANIS: Sean?

DR. TUNIS: Just a few questions about the tech assessment. One is, you mentioned at the beginning that you had excluded abstracts from review, which I know is a common thing to do. But can you say anything about the number of recent abstracts and the size of those studies and whether there is sort of a body of data about to emerge that's in abstract form now, or give us any feel for that body of literature?

MR. SAMSON: I did look through that body of literature, and I would say there is about eight or nine abstracts that haven't made it into print yet, and they cover a variety of uses. One is even on screening, which is slightly different from the indications on detection of breast cancer that we have looked at here in this assessment, but they don't add anything substantial to the assessment, they wouldn't change the conclusions.

And the primary reason for excluding them is that we just don't have enough information from them to be able to evaluate their methods and the quality

00071

of the study. But in terms of quantity of evidence, it's not a large body.

DR. TUNIS: Another question I had is, it seemed that the two key quality features of studies that were, essentially no studies that met these, were the verification bias issue and the blinding of, I forget which one it was.

MR. SAMSON: Of the reference standard to the PET result.

DR. TUNIS: Right. I'm just wondering, just for the nonmethodologists and the pseudomethodologists here, if you could just explain, you know, what is verification bias, how important is it, and the other as well?

MR. SAMSON: Well, in terms of weighing how important verification bias, that's a difficult thing to do. Methodologists are trying to come up with rating scales, but it's difficult to weight one form of bias against another, but it is agreed that it is an important source of bias. It happens when patients who undergo the PET scan have those results, those results influence the decision whether to undergo the reference standard. Now ideally you want all patients who get the index test to undergo the reference

00072

standard, so you don't want the results of the test to determine whether patients get the gold standard tests. It can bias the diagnostic performance data.

And the other question about whether the reference standard was blinded with respect to the PET imaging, we just couldn't find any studies in which there was a clear statement in the methods of the paper that that was done, and I can't explain why. I have seen it other literatures and diagnostic tests, but it just

didn't occur in this one.
DR. TUNIS: So the expected impact of that would be so the readers would, you're saying, may have known what the PET result was when they were reading the conventional imaging?
MR. SAMSON: We don't know, they could have, but we don't know.
DR. TUNIS: You just don't know.
MR. SAMSON: Yeah. The important point I want to make is that when I give those counts on the study quality characteristics, when I say zero studies had the reference standard interpreted blindly to the respective PET, the rest of them actually were just uncertain, we didn't have

00073

enough information to make a determination, but we couldn't say that any of them definitely used a blinded interpretation of the reference standard.
DR. GUYTON: David, the reference standard that you're talking about here is the pathology result?
MR. SAMSON: Well, it varied from indication to indication. For the first two indications, it was pathology, right.
DR. GUYTON: So, I don't see how knowing what the PET result, how that would affect the reading of the histology.
MR. SAMSON: You could make that argument, but it has also been argued in the literature that blinding of both reference standard and the test itself can have an impact on the diagnostic performance, and it has been studied to see if there is an impact and an impact has been found.
DR. MCNEIL: There is at least one article to look at the importance of verification bias, and it was by Colin Bage a number of years ago from Memorial Sloan-Kettering, and that looked at, I have forgotten, CT, and help me, and liver scans and liver metastasis, and the difference --

00074

it was an old study but it was actually quite a well-done study, and the difference in sensitivity among those who actually had the biopsy versus those whom they modeled would have had a result had they had the biopsy but didn't, which is a little bit tricky to do, but nonetheless, they did the best they could. It was quite substantial, I think it was about 20 percentage points in sensitivity, so that was a big hit on the verification policy. It increased it, the bias increased it.

DR. PAPATHEOFANIS: I had a couple questions for you, David. I notice that in your selection of evidence, you focused on papers that dealt exclusively with breast cancer. As this is an emerging technology, a lot of the literature includes compilations of different kinds of cancer in the same manuscript, where say a paper has 75 cases on breast cancer and maybe five on lung, I'm just curious, is there a significant number of papers that were excluded because of the purity of that criteria?

MR. SAMSON: I didn't keep a close count on that, but just my memory is no, there wasn't a lot of evidence that was excluded based

00075

on that. If we had included it, you would be mixing diagnostic performance data for PET in other malignancies and there could very well be different levels of diagnostic performance across different malignancies.

DR. PAPATHEOFANIS: Sure. My other question is, I guess also coming from a pseudomethodologist perspective. The confidence profile method for doing meta-analysis, David Eddy is a champion of that, and especially its use in diagnostics. Did you consider that? I know you used the random effects because of the Annals and the Cochrane approach, but did you consider using that?

MR. SAMSON: We didn't. We haven't accumulated much experience in using the

confidence profile method for diagnostic tests, and we decided to use an approach that has I think a better track record, at least it has been published on the summary ROC curve method, I think there is more literature on that and more people are familiar with it.

DR. PAPATHEOFANIS: Great. Any more questions? Michael.

MR. KLEIN: Yes. Did you consider in

00076

your recommendations or in your findings the comparison of PET to that of traditional film or analog based mammography where detection rates have been confirmed by a number of studies and average anywhere from 77 percent to 82 percent, that the false negative rate is in the high 15 to 20 percent rate and you have false positive rates also in the range of anywhere from 7 to 10 percent.

MR. SAMSON: Are you asking me this with regard to the screening use of PET?

MR. KLEIN: Correct.

MR. SAMSON: That was an indication that we were evaluating for this technology assessment. And basically, there are no data for using PET in the screening population. We just, that kind of information is not available so we didn't consider it.

MR. KLEIN: And then the other question I had, was the two to four centimeter range size selected, which would be indicative of a mid to early late stage cancer, was there a particular reason for that population chosen?

MR. SAMSON: That's just how the investigators selected their patients. The only

00077

guess I could make is that you know, whenever a diagnostic technology is being introduced, the investigators tend to test it out first on patients who have more easily detected disease. I can't think of any other reason why there is not

more data on patients who may have a lower prevalence of disease, maybe indeterminate mammograms and smaller tumors. I think it would be terrific if we could get that kind of data, but it's not available yet.

DR. PAPATHEOFANIS: Great. Any more questions? Well, we're three minutes behind schedule. Let me remind you that David will be around as he said all day, so we can return to him. I think you've done a fine job in bringing this data together for us and it has been very useful for us to have your document as we will go through the day. So let's take a break and return in 15 minutes.

(Recess.)

DR. PAPATHEOFANIS: We would like to get started again.

MS. ANDERSON: We are now going to move into the time for scheduled public comments. Public attendees who have contacted the executive

00078

secretary, that would be me, prior to the meeting will address the panel and present information relevant to the agenda. Speakers are asked to state whether or not they have any financial involvement with manufacturers of any products being discussed or with their competitors.

We are going to begin with Dr. Sam Gambhir, to be followed by Dr. F. David Rollo, Mr. Bob Britain, and Dr. Steven Larson to finish. Dr. Gambhir.

DR. GAMBHIR: Great. I'm actually over here, gentlemen, ladies. Since you're going to be looking at the screen, I figured I might as well stand over here, and she will operate the slides.

So in the 20 minutes that I have been allocated, I am going to use some strategies to try to convince you that what we're looking at is actually a different scenario than what's been presented during the last hour, hour and a half. I base this on going through the report that has been done and you heard presented, I base it on my

experience in actually reading PET scans over the last nine to ten years, and based on talking to the members of the oncology PET community. And I also add to this that because I build decision

00079

models myself, look at health care outcomes myself as part of a decision modeling laboratory, that I think put together a clinical picture with the health models in an appropriate way. Next slide.

So I will do this by first arguing that when we look at the breast FDG we cannot look at just FDG applications in the breast. I have argued this six months previously to the executive panel that we need to look at not just breast literature because in fact my belief is that, as other believe, that with PET, we are actually monitoring things that are -- can you guys hear me?

I will just speak up. I am going to argue for that briefly. Then I will take you through some literature reviews including some abstracts and tell you the importance of that. And I am going to argue for three areas for the use of FDG-PET. And the common theme in these three areas is that we need to look at the data we have available now, as limited as one might believe that data is, and look to see which women are the most underserved that can currently be helped given the understanding that we have of the literature. I will do this by simply looking at

00080

women with dense breasts, and then looking at women with recurrence, and finally going on to monitoring treatment and looking at FDG-PET in monitoring treatment, and finally I will conclude. Next slide.

Glucose metabolism and FDG are based on many many years, many decades of underlying biochemistry, well documented in the basic science literature. This has been stressed over and over again, but we need to remember that glucose

metabolism is critical to proper cell function, it's critical to cerebral function because of ATP derivation in the neurons, it's critical in ischemic tissue because it's protected, and in cancer specifically, it's increased 19 to 25 fold. I think when we look back 20 to 30 years from now, we will not look at cancers based on their site of origin, we will look at cancers based on their molecular errors, based on which alpha genes are amplified, which receptors are overexpressed or underexpressed.

And really that alludes to the fact that the literature you look at for FDG needs to look at all cancers, not just breast cancer. What causes a false positive in breast cancer in many

00081

ways is similar to the locations in which that lesion is found, and that can be similar to lung cancer, similar to a head and neck cancer that's metastasized. It's not just the origin of the tissue, it's the common need for glucose. Next slide.

This of course goes back to the biochemistry of these cells needing to produce ATP through their high proliferate rates. In fact using anaerobic glycolysis, less ATP per glucose is produced so more glucose is needed. Up regulation of glucose transporters and hexocynase then drives the various pathways for both energy derivation and DNA and RNA synthesis. These up regulations are common to breast cancer cells as well as lung cancer cells, as well as a whole host of other cancers, and breast cancer cells for the most part are on the high end of the spectrum, they are not on the low end of the spectrum in terms of up regulation of fundamental molecular pathways. They tend to take up a lot of glucose and therefore, a lot of FDG. Next slide.

Now we have heard extensively the literature review, which I've also reviewed independently, and I have no disagreements with

00082

it. The studies are limited, there are needs for improving those studies, there's reasons to improve them, I think that will happen in due time. Next slide.

But what I have done is just illustrate a few, and when you cut through all those studies that were presented, whether they be research articles or more recently abstracts, what we're dealing with is, yes, a handful of articles. But they are being published not just in imaging journals but in cancer journals like the Journal of National Cancer Institute, Journal of Clinical Oncology, surgical journals as well. They are not a series of limited articles in limited journals. And yes, each of these do have limitations, they could have larger numbers, but this is again that catch 22 that without reimbursement it's very difficult to do the larger kind of studies that need to be done because these are of course not being backed by any sort of clinical trials from drug companies or manufacturing companies. Next slide.

These studies do date back all the way to even 1989. They have slowly built up to the most current year where I think we will see a

00083

whole host of other studies. I reviewed myself five papers that are currently in press that are not available to anyone, three of which will appear in the Journal of Clinical Oncology. All of them continue to point to building evidence based on the kinds of preliminary data that these earlier studies generated. Next slide.

When you break down based on each of the categories, diagnosis, staging, recurrence, as well as monitoring treatment, and you look at articles and abstracts as well as articles only, for the most part, including the abstracts strengthens the end, it increases the total number of patient studies, and it tends to actually decrease slightly the sensitivity and specificity.

That is, the abstracts I think are showing us that the actual accuracies are dropping slightly compared to what we saw in the research articles alone, but in fact gives us more weight that these actual sensitivities and specificities are reasonable.

For example, in the area of diagnosis, what we're looking at are sensitivities of 90 percent and specificities of 92 percent, with an overall accuracy of 88 percent, when you look at

00084

lesions. And when you come to the research articles, those are the only ones that also looked at lesions so it's the same numbers, 90, 92 and 88. But when you look at patient studies, if you just look at research articles, we've gone from about a hundred to 200, double the number of patients. The accuracy when you look at research articles alone are 93 percent, 93 percent, overall accuracy of 94, where here we're looking at 91, 93, 95. Not much of a change, even though we've doubled the number of patient.

Now these admittedly are based on these weighted averages, these are not ROC analyses, but as I will argue, I don't think the real issue is what is the exact sensitivity or specificity. It's a range of sensitivity and specificity that continues to be reinforced based on the outcoming data. Next slide.

In staging, what we're seeing again when we include just articles are sensitivity of 92, specificity of 90. When you include articles and abstracts, again, a significant jump, about 500 more patients. What you see are accuracies of 91, 88, and overall accuracy of 90. So the values are not changing that much, although we're gaining

00085

more confidence that these results are real based on the large number of patients. Next slide.

When you look at diagnosis and staging combined, these are just a very limited subset of

data, and again, there is no significant addition through the abstracts. Next slide.

When you look at recurrence, again, doubling the number of studies from about 200 to 400, you see sensitivities of 90, 90, 93, and 80, 85, 82, so slight decreases in the sensitivity and specificity and accuracy based on adding the additional number of studies. Next slide.

In monitoring response, again, what we see happening is going from 150 to about 200 studies. Sensitivity of 81, specificity of 97, accuracy of 92, and now we're going to 81, 96, 92, so again, a very similar pattern. Next slide.

So what these data are telling me really is that if you just look at research articles alone, after you dissect apart all the different areas of applications, in over about a thousand patients right now, across just the research articles, what we're looking at are ranges of sensitivity of 75 to 91, and specificity of 74 to 93. When you add in the abstracts, these

00086

same kinds of ranges still persist; as a matter of fact, the ranges don't change, the abstracts all fall in between these ranges, it's just that the number of articles and abstracts now increase, and the number of patients comes up into the 2,000 range. I don't think the real issue, even if we were to revisit this problem two or three years from now, will be what is the exact sensitivity and specificity of PET in breast cancer detection, diagnosis, management, recurrence. We can continue to gather the studies and my best guesstimate at the current time is they will continue to fall in these ranges.

I think the bigger problem is, what is the clinical applications in which women that are currently underserved would benefit from a sensitivity and specificity in the current range as compared to what other studies offer us, and that's how I have modeled the next set of arguments. It's not going to be about trying to

find out for sure what the exact sensitivity and specificity are. What we're really going to have to ask is what studies are patients going to benefit from the most and in what clinical management outlooks. Next slide.

00087

So, that's what we want to focus on, what's the clinical setting based on what we know about accuracy today so we can ration the technology for a good use. Next slide.

I think one such application that we have not addressed properly in the literature but there is inference based evidence for to apply PET to is that of women with dense breasts. Next slide.

This is an example of a female in her young 40s who actually has a high risk of breast cancer based on family history, who kept getting mammography even after the age of 35, kept having negative mammograms, had dense breasts, Wolf scale DY based on mammographic density, and then finally because she had access to it, decided to have a PET scan based on being able to pay for it herself. And there's clearly as it turns out, to be a one centimeter focus. This is ductal, infiltrating ductal carcinoma. This is an example of the kind of signal you can get from a dense breast. This signal is compromised in low energy x-rays of mammography. PET has no real concern that this is coming from a dense breast. The physics are such that that signal is properly

00088

relayed from a dense breast.

In fact, you have to focus not just on that, but actually in other views where her arms are up, lack of axillary findings and lack of involvement in her entire body. What would happen to her had she not have had this PET scan? Well, no one can say for sure, but my guess is this lesion would have continued to have gotten bigger, bigger than the one centimeter it is, eventually

it would have been palpated, would have been found, and she would have had a chance for staged progression. Next slide.

In fact, four years ago, I started looking at, with Matt Allen in my laboratory, decision models for just this area of application, not specifically for PET, let's try and understand, what can we do for women with dense breasts that have a mammogram, are actually falsely negative, and then just come back and have a screening mammogram a year later and a year later and always miss, what can we do for these women now given that they are not being well served by the existing modalities. Next slide.

Well, one possibility is in fact concern for these women that have a negative

00089

mammogram, a PET study. We don't have data specifically looking at the use of PET in this exact population, but we know based on the physics of the technology, based on applying it to women with dense and nondense breasts together, that PET has a sensitivity in the range that I'll show you an a specificity in the range that I'll show you. And these women have no other way of knowing really whether they have a focus. Next slide.

Does it matter that you catch the dense breast lesion early? Yes. I think at this date in research and treatment, that there is evidence to believe that for a six-month delay or more, women would have a .047 chance of distant disease on initial presentation. That's with a six-month delay in diagnosis, but when there is no delay, that is when they're caught right away, and of course we catch more at the local stage, .875, less at the regional, and none at the distant. So this six-month delay is costing a progression of disease in these women that have in this case dense breasts. These women then will benefit potentially if we can insert a test to catch them prior to stage progression. Next slide.

In fact, mammography does terribly in

00090

this area. Just like that woman who I showed you, the example where the sensitivity was unknown but the mammogram missed the lesion for three years, 66 percent is the estimated sensitivity for mammography in dense breasts based on the literature. We have done sensitivity analysis to cover the entire range. PET, even if you go to the lower end of the range, you would be looking at 75 to 80, maybe 70 percent in this range. For these models I've actually plugged in 70, but I'm showing you that realistically I believe it is higher, that it's around 80 percent sensitivity.

The specificities of mammography and PET are comparable in this application. The biopsy approaches are not accurate, at least for a needle biopsy. Incisional biopsy essentially is 100 percent. We've modeled all these. The details are in that article in Breast Cancer Research and Treatment by Allen, et al. Next slide.

What we've shown is that you if you screen, for example, 3 million women with DY breast density with mammography and a second test, and in this case I have inserted an FDG-PET, that will lead to 1,638 fewer false negatives than

00091

using mammography alone. That will translate to a prevention of 267 women progressing from local to regional, and 78 women from regional to distant disease. These numbers don't sound big when you compare it to the number of women being put into the algorithm, but remember, the incidence of breast cancer is pretty low.

If you want to be even more selective about who you choose, rather than just DY breast density on the Wolf scale, of course you can reduce the number of women coming in. But these women are underserved, they are going to have their lesions missed on mammography, they are going to progress in stage, and these studies as

the best I can do based on the available literature, reasoning from what we've got, show that we're going to actually help to prevent progression of the disease, so I would like for you to consider that as one potential area of application, with an actual health outcome difference, not just a simple, you know, here's how many scans PET avoided or here's how many costs it saved, it's actually a health outcome difference. Since we're looking at health outcome, let's go after women that are

00092

underserved. Next slide.

Then how about assessment of extent of disease after recurrence? Next slide.

Here is an example of a woman who had breast cancer actually in the right breast, which is on our left, had lumpectomy, had adjuvant chemotherapy, three or four years later presents back with rising tumor markers, in this case a mucin marker, and is now subjected to a whole body PET scan. Again, the power of PET here as an initial tool is that the entire body is surveyed, we can immediately get a sense for where the disease may be localized. In this case, it's actual in lymph nodes, in the mammary lymph node and sternal bony involvement. More importantly, there is not any involvement in the axilla, the breast mass, as well as distant metastases in the abdomen or pelvis. This directly influences management because now she can undergo locoregional treatment as opposed to more systemic treatment, although some would argue that her bony metastasis of the sternum would dictate more aggressive treatment.

Does it make a difference in terms of health outcome? Not clear, and I can't argue

00093

these issues in terms of health outcome, but I can argue them in terms of at day-to-day practice. Physicians will routinely after, whether a tumor

marker triggers recurrence, a new palpable mass, a new palpable node, start by doing a series of CT studies, start by doing a series of bone scans, try to see where the tumor has recurred, has it recurred regionally, has it recurred regionally plus the axilla or distant? And the key is, PET is giving you all that information in one scan, yes, with not a perfect specificity or perfect sensitivity, but with, as you heard, close to a specificity and sensitivity with what we see in the other modalities.

And for this you can look beyond the breast literature. What causes the false positives and false negatives in the abdomen and pelvis for breast cancer are the same as what is the case in lung cancer, it's the same underlying biochemistry for these lesions, it's not new just because it originated in breasts. Next slide.

Here's another example where pleural metastases now have occurred, dictating a much poorer prognosis. This is someone who actually had left-sided breast cancer, on our right, and

00094

again, three or four years later started to present with in this case rising tumor markers, CEA actually, and was found to have excessive pleural involvement. Next slide.

So, in 35 research articles, the mean sensitivity and specificity are around 90 percent, 90 percent sensitivity, 90 percent specificity, and including the abstracts, there's about a 40 percent change in management occurring on a day-to-day clinical practice. Use of FDG-PET in this setting would help to establish the aggressiveness and the nature of treatment. Yes, I don't know if it will make a difference in health outcome for these women, but it will in fact establish the nature of the chemotherapy or local aggressiveness if you wanted a local surgery, based on the FDG-PET study.

So again, these are women that would be better managed if the extent of disease throughout

the body could be better identified and more appropriate management undertaken. Next slide.

Monitoring response to therapy. Next slide.

Here's an example of the kind of typical studies we see. I think one injustice in

00095

looking at these articles and the numbers is that you don't get to see the visual results. This is a woman that has in fact recurred, there is -- I'm sorry, was initially diagnosed with Stage Iv with extensive nodal involvement and metastases. You can see FDG uptake in lymph nodes and in bone throughout the thorax. Within two cycles of chemo for her you can see resolution of those same foci, well in advance of the CT which still shows no potential size reduction. These women can be better managed because in fact now in her case, we know the chemo is working.

Similarly, there are example where the chemo is not working and we can change the therapeutic options for that patient. Next slide.

So yes, we only have a few research articles, we have five research articles. The mean sensitivity and specificity are 90 and 74 patient, and it's only 174 patients so far. But again, I think as more data will be generated, these sensitivities and specificities will continue to fall in these ranges and really the case will be what is the management change based on these studies that are occurring. And again, the data continues to show in clinical practice

00096

that FDG-PET in this setting would help to establish response or lack thereof for a given treatment regimen, allowing you to change the chemotherapeutic regimen and this should lead to better management, I don't know if it will lead to better health outcomes, but on a day-to-day practice that will lead to better management and hopefully earlier response can be gauged and

hopefully better outcomes will result. Next slide.

Some conclusions. FDG-PET has a biochemical basis that will continue to reinforce the accuracy of this test in various clinical settings. Please do not look at just the breast cancer literature alone, that's not the right way to think about malignancies. The mean sensitivity and specificity values for FDG-PET in various applications for breast cancer are not likely to change much with the additional studies. Yes, they may fluctuate around these different means and variances, but they are not going to change the overall value significantly.

It's underserved women, screening women with dense breasts, restaging women with recurrence so we get a whole body survey,

00097

monitoring response to therapy so we can change chemotherapy are all applications that can be currently justified if you look at management changes, some minor decision modeling with the evidence we have so far. Next slide.

We need to focus FDG-PET on women that are most underserved. This will allow and justify the rationing of the technology. We don't want to use it unnecessarily across every indication but the indications I'm showing you I truly believe have practical implications, have enough data to give us some jump in trying to study these women now, and we shouldn't wait to help women in the future when we have ongoing validation, rapidly emerging abstracts that enforce the data, when we can help these women now. Thank you.

DR. PAPATHEOFANIS: Great, thank you, Dr. Gambhir. I wanted to take just a couple minutes to see if there are any questions for Dr. Gambhir by panel members before we go on to the next speaker. No? Okay. Let's go on to the next speaker. Dr. Rollo.

DR. ROLLO: Good morning, I am Dr. David Rollo. I am currently the chief medical

officer of ADAC Labs, a company that was recently

00098

acquired by Philips Medical Systems. ADAC Laboratories is a manufacturer of PET imaging systems. I joined ADAC in October of 1999. In this position I am responsible for the clinical research programs, luminary and professional relations in the management of the medical advisory board. I am also the medical director for all regulatory compliance matters and serve on the strategic planning committee of the corporation.

Previously I was chief medical officer of Humana when Humana was a hospital company and also at that time was also the owner of its own medical insurance plan. At Humana I held the position of senior vice president of medical affairs as well as the founding medical director of the Humana Health Plans.

In addition, I am on the board of directors of the diagnostic imaging and therapy systems of the National Electrical Manufacturing Association, known as NEMA. I am here this morning representing the views of NEMA.

NEMA is the nation's largest trade association representing the United States' electrical industry. NEMA's diagnostic imaging

00099

and therapeutic systems division represents more than 95 percent of manufacturers in a $5 billion market for high tech x-ray imaging, computer tomography, diagnostic ultrasound, radiation therapy, magnetic resonance imaging, and nuclear imaging equipment. In addition, the division represents the manufacturers of picture archiving and communications systems.

I am accompanied this morning by Mr. Robert Britain. Mr. Britain is vice president, medical products, of NEMA. Prior to joining NEMA in 1985, Mr. Britain spent 23 years in the United States Public Health Service Food

and Drug Administration, during which he held positions as Director, Office of Compliance, Bureau of Radiological Health, Deputy Director of Bureau of Medical Devices, and Director, Office of Device Evaluation in the Bureau of Medical Devices and the Center for Devices and Radiologic Health. Mr. Britain is here to assist in the event that any policy issues relating to the medical imaging industry be raised.

On behalf of NEMA and its member companies, we appreciate the opportunity to address the Medicare Coverage Advisory Committee

00100

for Diagnostic Imaging Panel on this important topic of Medicare coverage of breast cancer. There are a number of issues, concerns and considerations that we would like to urge the panel to bear in mind as you deliberate on this coverage issue.

First, we'd like to urge the panel to consider a wide body of evidence, move forward to consider PET coverage for breast cancer. No one on this panel needs to be told of the devastating effect that breast cancer is having on women and their families across this country. The stakes are huge. According to the American Cancer Society, every woman is at risk for breast cancer, and as a woman ages the risk of breast cancer increases. This year 182,800 women in the United States will be diagnosed with breast cancer and over 40,000 of these individuals will die. Excluding skin cancer, breast cancer is the most common form of cancer in women in the United States and is the third leading cause of cancer related deaths. We believe the breadth and scope of this deadly disease requires a flexible and forward looking approach to providing not only new tools in the diagnosis and treatment of the

00101

disease, but also to encourage an environment that is conducive to the development of new

technologies to address and eradicate this disease.

The key to successful breast cancer treatment is early detection, finding, accurately staging and treating the cancer before it has had a chance to spread. The five-year survival rate for localized tumors, that is tumors that have not spread out of the breast tissue, is nearly 97 percent. For those that have spread to adjacent lymph nodes, it is around 75 percent, and for cancers that have spread to other parts of the body, it's only 20 percent.

Clearly technology advances will be able to better whether suspicious structures are in fact malignant and whether or not any of the malignant cells have metastasized to adjacent or distant parts of the body. Clearly one of the real remaining challenges of diagnosis and treatment for good breast cancer is good staging. As everyone here knows, PET is a noninvasive diagnostic procedure that assesses the level of metabolic active and perfusion in various organ systems of the human body. In PET, the positron

00102

camera is used to produce cross-sectional tomographic images which are obtained by imaging a positron emitting radioactive tracer such as FDG or fluorodeoxyglucose. This is usually administered intravenously to the patient.

This technology has proven valuable in providing metabolic information on tumor activity and other indications. Currently, HCFA is covering PET for diagnosis, initial staging and restaging of non-small cell lung cancer. For colorectal cancer, it has been a standard to include diagnosis, staging and restaging. It's also the initial staging and restaging of both Hodgkin's and non-Hodgkin's diseases, the diagnosis, initial staging and restaging of melanoma, the diagnosis, initial staging and restaging of esophageal cancer, and the head and neck cancers.

And importantly, as we just pointed out, in all cases we are looking at increased metabolic activity having nothing to do with the source of the cancer, but simply the process of what happens when cancer cells spread to other parts of the body.

Congress and administrations past and

00103

present have recognized the importance of moving forward on the diagnosis and treatment of breast cancer. Federal spending for breast cancer research at the Department of Defense, the National Cancer Institute and other federal agencies has grown over the years, and it has become quite significant. This reflects the intense interest in the mind of the public in bringing the resources of the federal government to bear on saving the lives of the women who are diagnosed with this deadly disease.

We believe that this context should drive this panel, and subsequently the full MCAC Executive Committee, to take into consideration the full array of clinical information about PET's effectiveness, such as experience of practicing physicians, medical specialty societies and patients. We do not believe that the analysis of evidence about a technology's effectiveness, especially in dealing with a deadly disease such as bread cancer, should be confined to peer reviewed articles, which are the sole source of information for the Blue Cross/Blue Shield Technology Evaluation Center report. We believe it is appropriate for HCFA and for you in your

00104

advice to this Agency to take into account broader public policy considerations and coverage decisions.

For this reason, we believe it is both warranted and appropriate for this panel to take into consideration in addition to peer reviewed studies, the expert judgment of the leading

developers and innovators of PET technology, input from the appropriate medical societies, from patients, and the fact that the United States Government has made the improved diagnosis and treatment of breast cancer a national priority. For this reason, we also believe that the fact that PET has been determined by HCFA to be worthy of coverage for six other cancer indications to be suggestive of its potential effectiveness in other indications should be recognized for breast as well.

Second, we are concerned that the panel consider the fact that in many cases, medical practice and technology evolve more rapidly than the publication of studies which document their benefit to patients, as we just noted. Technology assessments relying on peer reviewed published literature which meets preestablished rigorous

00105

inclusion type criteria such as the TEC assessment do not adequately and fully reflect the current practice of medicine, or available technology advances for patients in existence today. Timeliness in coverage decision making is essential to providing access to patients to the latest innovations of medical technology.

Third, we believe HCFA and the MCAC should explore ways to insure that Medicare beneficiaries have access to emerging medical technologies, not just existing an mature technologies, while at the same time providing for development of information to support decision making in the long term. Clinical experience and actual patient studies should be considered, along with patient registries, real-time data collection, or collaborative agreements with other bodies as possible alternatives.

For these reasons and for this indication, we believe that the panel should have and should exercise reasonable flexibility, again, the word flexibility, in its coverage recommendation. This flexibility should extend

not only from the nature of the evidence considered for the effective of PET for breast

00106

indications, but also to a forward looking coverage policy for a national priority disease that lays the foundation for subsequent studies and data collection that would support longer term coverage decisions.

Fourth, we believe there are important indications that PET has a unique capability in terms of its value in staging breast cancer and detection of metastatic disease. In comparison with other diagnostic modalities, PET possesses a greater degree of sensitivity and specificity that enables it to detect metastasis far earlier in the disease process that permits appropriate and timely treatment of metastatic as well as localized disease.

The clinical information obtained from PET imaging can be used to avoid or sharply reduce the cost and risks associated with surgery on patients with inoperable cancer, which is also a consideration for the other indications that have already been approved.

My personal experience is at Cedar Sinai Medical Center in Los Angeles, where we are conducting a clinical trial on breast cancer. As an example, I recently participated in a study on

00107

a 35 year old woman with evidence from a mammogram and biopsy that she had a solitary cancer in her left breast. She requested a PET scan, with the understanding that she would have to pay for this at her own expense, because she wanted to be sure of the diagnosis, that is, she had no evidence of disease other than the solitary mass that had been identified. The PET study showed four lesions in her breast rather than one, as had been indicated on the palpable mass, the mammogram and on the biopsy, as well as lymph node involvement that was not evident on the examination by her referring

physician. The whole body study showed no evidence of additional metastasis.

Her treatment was changed from a lumpectomy, which had been the original decision by the surgeon and her referring physician, to a radical mastectomy and lymph node dissection. The staging clearly saved this woman's life and the agony of disease when it was detected months later from the residual if it had not in fact been removed. The cost implications would have been roughly $10,000 for the lumpectomy treatment, followed by a 60 to $80,000 dollar chase of the cancer that was left in her body, with a life

00108

expectancy over two years of no more than 10 to 20 percent. The staging using PET resulted in a 15 to $20,000 treatment for the radical mastectomy, lymph node dissection and associated chemotherapy. The treatment provided this 35 year old woman with an 85 percent survival rate at five years.

More generally as this experience confirms, we believe that metastatic staging using PET has the potential to detect distant metastasis in the liver, the skeleton and distant nodes. Importantly, the presence of distant metastasis radically changes the treatment from aggressive to palliative. Likewise, a patient's prognosis changes from hopeful to very poor. The PET survey potentially can replace the need to perform a conventional metastatic survey, including CT, ultrasound, and conventional bone scan. This approach could be especially valuable for patients with Stage III breast cancer at the time of initial diagnosis or in patients with suspected recurrence.

NEMA is aware that the use of PET for primary diagnosis of all breast lesions and staging for nodal involvement, while reported in the literature, is not reported for a

00109

statistically significant patient population.

However, we do understand that the reports that are in the literature, whether they be abstracts or full-blown articles, are extremely positive on the clinical value and the promise.

Sensitivities of 90 percent and specificities of 95 percent are reported by the Academy of Molecular Imaging as the values that could have a negative predictive value of 97 percent, and spare 33 patients the morbidity associated with the axillary lymph node dissection at a cost of missing one patient with lymph node involvement.

Such high sensitivities and specificities have been reported with attenuation corrected studies, which were not reported this more or at least were not segmented out from the literature. The lit number, however, indicate clearly numbers in the 75 to 90 percent range for both sensitivity and specificity for mainly nonattenuation corrected images.

Finally, continued availability of technological advances for patients depends on the ability of medical device companies to devote research funding for their development. An

00110

environment that is conducive to the steady flow of new medical technologies to address the health needs of the American public should be a concern of the federal government. Coverage and reimbursement decisions made by HCFA have a critical and direct impact on the ability of companies to dedicate funding for research and development, and adverse decisions could have a negative impact on the development of new technologies.

For the past two years, the trade association I am here representing today, NEMA, has partnered with the National Cancer Institute to hold an annual symposium in Washington D.C. designated to facilitate communications between industry, academia and the federal government in order to stimulate further research and

breakthroughs in medical imaging technology. There is little doubt through these symposia that the real excitement and hope for breakthroughs in the imaging field are on the area of molecular, gene, and other biomolecular imaging modalities, of which PET is considered the vanguard.

The hope is for technologies that not only improve diagnosis and sharpen our range of

00111

therapies, but ultimately for technologies that will enable us to image therapeutic interventions in real time at the molecular and gene level in order to evaluate the effectiveness of a given treatment regimen.

If this sounds exciting, it is, but the development of these technologies is not a foregone conclusion. One of the things that industry has learned from these symposia is that we need to do a better job of education of our friends in academia as well as government in how companies make their investment decisions. Research and development funding in most companies must vie against many competing interests inside the company, marketing, operations, expansion, capital equipment, acquisitions, current product enhancement. In most medical technology oriented companies there are many research ideas, far more than can be funded by the dollars that are available.

Vice presidents and directors of research and development at medical technologies across this country are forced to make difficult decisions on what projects they will fund each year. Many factors go into the decision between

00112

the winners and losers in this process. Some of these factors include clinical need, the reimbursement climate, the risk of the project, the overall cost of the project, time to market for any new products, as well as potential size of the market.

The strength of your intellectual capital is another major factor in determining where the funding will in fact be administered. Just because the technology is exciting or potentially revolutionary does not mean that most medical technologies are going to invest the time and the money to develop it, especially if the market is small or difficult to enter; companies will either keep their investments modest or make no investment at all. In the field of medical technology, one of the key considerations in determining the size and difficulty in entering a market are medical insurance coverage and reimbursement decisions. If it is expected or proved difficult to gain coverage for new technology, or if reimbursement levels are such that there are few incentives for providers to purchase the equipment, there will be no strong pressure for companies to place their R&D dollars

00113

in this technology, no matter how exciting the promise of the new technology.

With PET set on what appears to be an indication by indication coverage path, there is no doubt that this is a difficult market to enter and the prospects for recouping R&D dollars may be long and arduous. In this context, while we recognize that this panel's responsibility is to make recommendations with regard to coverage for a given technology, for a technology so widely thought to be promising and especially for an indication whose diagnosis and treatment are a national priority, we believe it is appropriate and justified for you to exercise flexibility in considering this decision coverage, and to factor into your considerations the potential impact a negative recommendation could have on future company based R&D in this promising technology field.

We appreciate the opportunity to raise these issues before you today and would be pleased to answer any questions you might have. Thank

you.

DR. PAPATHEOFANIS: Thank you, Dr. Rollo. Any questions from the committee at this point?

DR. TUNIS: Maybe just one question. I guess, just so I understand a major aspect of your discussion is obviously the priority and importance of breast cancer as a disease entity in terms of morbidity and mortality, et cetera, which I think no one disagrees with. And your asking for a flexible approach presumably reflects what's in the TEC assessment and what Dr. Gambhir mentioned, which is that the quality of the evidence is sort of acknowledged for most of the proposed clinical applications to be weak, I presume that's what a flexible approach leans toward.

I guess the question I would pose to you is, why wouldn't that approach lead to this panel recommending coverage for primary screening for breast cancer using PET as opposed to conventional mammography? In other words, what is this panel supposed to look at to answer that question in the negative and these other questions in the affirmative, or how would you parse that so that the panel could sort of think through where they might draw the line in terms of applications of PET?

DR. ROLLO: Okay. The answer to that, in terms of flexibility, we're looking at breast imaging in much the same way that HCFA approved the indications for other cancers. It literally was on a trial basis, the promise that in fact the staging of patients with other cancers could in fact lead to more appropriate treatment and management, and by that we mean the elimination of surgery that may not in fact be beneficial to the patient, the elimination of many heroic procedures that physicians could administer to patients as

opposed to palliative treatment, knowing full well that the extent of the disease based on statistics would indicate that that patient had less than some particular time to live.

And rather than having them go through that agony of that type of treatment -- let me give you an example of what I'm saying even more specifically. When I was at Humana, one of the things I developed is 90 dedicated breast clinics. These were clinics that were dedicated to women, they were separate units within hospitals, separate from the diagnostic radiation department. The focus was on education, it was on self examination and it was on the evaluation and

00116

encouraging patients to have screening examinations. What we found is that we found a lot of cancer that probably would not have been detected otherwise in part because we also encouraged payers, for the individual employees who had a plan, to offer free screening to their employees.

What we found, though, is that when they made the diagnosis, invariably the physicians were looking at lumpectomies as the alternative. If it was a solitary mass, they would do a lumpectomy. We had literally hundreds of cases where the lumpectomy in fact failed to reveal the fact that the patient had either axillary nodal involvement or additional lesions within the breast, and we ended up with huge expenses. The numbers I gave earlier are Humana numbers on what it cost us to chase the disease.

By doing the staging and doing the procedure that we're talking about with PET, using a flexible approach, it doesn't say we're going to do everything, all the four indications that were suggested before, but looking at this as a better way of making a diagnosis. And part of it would be, the mammography still ought to be the

00117

screening procedure. The biopsy still should be the standard in defining whether or not that lesion happens to be a cancer. But the next step would be to use the PET imaging to determine whether or not there are additional lesions within that breast, nodal involvement, and also metastatic disease throughout the body.

So it would be kind of a staged approach with many of the types of data that Dr. Gambhir had suggested, and we know that the literature is beginning to prove that this in fact would be an approach that would be worthwhile. So we're looking at flexibility, not giving it all, but really looking at it as something that would be developing.

DR. MCNEIL: David, just following up a little bit on Sean's question. I'm a little confused about where we draw lines in terms of what's within our decision portfolio today. And as a follow-up question, I would like to ask you the following: Would you make the same argument that you're making now about, we're talking about screening for example, let's just take Sean's question about screening women, or even looking at axillary nodes for women, just taking the distant

00118

disease out of it for the moment. Would you make your same argument about MRI and would you think therefore, that this analysis should focus on a comparison between PET and MRI, for either the main indication number one, or the subset of that that Sam mentioned in terms of dense breasts? What I'm trying to do is really parse what we're deciding because this is really an overwhelming field, and it's a little bit hard for me to figure out where we make decisions and with what database we use them. I'm a little concerned about using just anecdotes, so can you help me through that process?

DR. ROLLO: I think as we all know, there is a dramatic information in terms of the information gained from MRI as opposed to PET

imaging for looking at early detection of metastatic disease.

DR. MCNEIL: No, I'm talking about not metastatic, local.

DR. ROLLO: Okay. To me, I thought the question had to do with, would we make the same argument for MRI as a screening.

DR. MCNEIL: For screening, right, for indication number one.

00119

DR. ROLLO: And I'm not thinking of this as screening as much as I am for staging of the cancer. Once the cancer diagnosis has been made, just as we did in lung cancer and other indications that have already been approved initially, it was not for diagnosis, it was not for screening, but rather for evaluation of the presence of distant disease for purposes of determining the most appropriate treatment and management for that particular patient.

So I'm not thinking of this in the sense that people were suggesting that if we have a palpable mass we immediately go to PET as a screening procedure to look at or eliminate the need for biopsy. I'm looking at it strictly as a staging, just as we did in the original indications for PET imaging, once we've got the diagnosis to determine the extent of the disease.

DR. MCNEIL: So you would not support its use in indication number one, is that what I infer?

DR. ROLLO: That's correct, right.

DR. PAPATHEOFANIS: Thank you, Dr. Rollo. Also, thank you to Mr. Britain for attending this meeting.

00120

I would like to call Dr. Larson up for his comments at this point. Welcome.

DR. LARSON: Thank you. I am Dr. Steven Larson. I am the chief of nuclear medicine at Memorial Sloan-Kettering Cancer

Center, and I have worked with PET now for over 20 years. My laboratory was one of the first to recognize that the altered metabolism of malignancy could be used as a basis for PET in the late '70s, and I started with PET development in the early '80s at the University of Washington, became the head of nuclear medicine at NIH in 1983, and then subsequently in 1988 -- in 1983 we developed a major PET program of which cancer was a fledgling development, but began to develop them, and then in 1988 went to the Memorial Sloan-Kettering Cancer Center, and since that time we have been actively developing PET in collaboration with our clinical colleagues.

Now, today I am representing as a member, the American Society of Clinical Oncology at the request of Dr. Larry Norton, who is the current president. And so, I would like to read you a statement from the American Society of Clinical Oncology.

00121

The American Society of Clinical Oncology -- this is regarding FDG Positron Emission Tomography imaging for breast cancer diagnosis and staging. The American Society of Clinical Oncology (ASCO) is pleased to have the opportunity to comment on FDG Positron Emission Tomography imaging for breast cancer diagnosis and staging. ASCO represents more than 16,000 physicians and health care professionals from 95 countries involved in cancer research and treatment.

Based on a review of the literature and other available evidence, we believe that the data support the following indications for PET-FDG scanning in breast cancer: PET-FDG should be used for imaging of suspected recurrent breast cancer, staging of locally advanced disease prior to therapy, and for monitoring treatment response in advanced breast cancer.

We would like to present additional data for the committee's consideration which is

based on a retrospective review of 133 -- and I'm sorry, there is a typo in this, it should be 133 patients with breast cancer who were referred for PET scanning at Memorial Sloan-Kettering Cancer

00122

Center in New York. We believe the data will support the indications for the use of PET-FDG scanning in breast cancer for recurrent cancer, locally advanced primary tumors.

DR. GUYTON: Dr. Larson, are we supposed to have this.

DR. LARSON: Yes, I believe you do have that in the packet.

DR. GUYTON: All right, I have that, but I don't have what you're reading, which is different.

DR. LARSON: I'm sorry, I thought we did provide it to the panel. I apologize. We can get copies of this for you. This is the letter which is on the ASCO letterhead for the committee.

So to continue, the data presented will support the indications for the use of PET-FDG scanning in breast cancer for recurrent cancer, locally advanced primary tumors, and for monitoring the treatment response in advanced breast cancer.

Thank you for the opportunity to submit ASCO's views on PET scanning for breast cancer diagnosis and staging. And I'm speaking on behalf of oncologists and their patients, and

00123

particularly on behalf of Dr. Larry Norton, and we urged the HCFA administration to consider covering this important procedure for those indications.

Now, if the chair, with the indulgence of the chair, I'd like to just talk a little about one of the abstracts that will be presented at the Society of Nuclear Medicine this year, which deals with FDG-PET scanning and the experience at Memorial. I think this goes to the point of providing many forms of evidence to the panel that

include the available evidence that we have.

What we did, I think specifically, if you look in the statement that's prepared for the Society of Nuclear Medicine, you see that it is one of the abstracts that's listed in the oncology tract, number 1236, talking about the impact of FDG-PET scanning on the management of 133 breast cancer patients. I think that this goes to the issues in our questions, especially question number four and five, which the committee has posed, namely looking at the more advanced disease.

Now what essentially this is, and I refer now to this little packet of handout materials that Dr. Guyton referred to, basically a

00124

review of the experience that we had at Memorial Sloan-Kettering Cancer Center, so this is the actual experience that we had in patients who were referred for PET-FDG scanning over the interval from May 1996 to July 2000. These are consecutive patients, they are the experience that we have, and so we looked at this to see if our experience with PET-FDG in our own patients referred by physicians for developing answers to clinical management issues, where that was consistent with published literature.

So then in 133 patients, and again, if you turn to purpose of this study, it was to determine whether PET scans affected disease outcome of breast cancer patients. And I will, it's a rather broad definition of that term. Again, in terms of materials and methods and study design, it's a retrospective study, so it has all those limitations. It is, however, as we have said, a consecutive review of all the patients during that interval who were referred by our clinicians for PET scans.

It was done with the most advanced available equipment that we have at this point, although equipment is evolving rapidly, as

00125

Dr. Rollo mentioned. We did this with a GE advanced dedicated whole body PET scanner.

One of the difficulties I think in this whole field is defining a gold standard, because obviously, one biopsy which may take a milligram of tissue in a person who is 70 kilograms, in some cases more than 70 kilograms, such as me for example, is really, there's all kinds of problems inherent in that. So one type of gold standard is just to use all the available evidence and to follow the patient for six months, and that's what we did.

So the confirmation of cancer was based on biopsy, correlative imaging which showed progression or stability, follow-up of clinical data and so forth. And then the clinical data was assessed at the end of a six-month period by informed clinicians to determine whether the patient's condition had improved or worsened under the treatment, and also what the impact of PET was on choosing that treatment.

The characteristics of the study patient are listed in the next slide. You can see that the majority were infiltrating ductal carcinomas originally and they were of a variety

00126

of stages, but the largest group was advanced patients. So these are advanced patients.

In terms of the characteristics of the study, the indications for the PET scan were conventional studies were equivocal, a frequent problem in advanced patients, especially after a lot of treatment has altered the appearance of more conventional techniques and when normal tissue has also responded to those techniques, such as radiation and surgery.

33 were referred for staging and restaging. This clinical suspicion of recurrence may have been an enlarging mass but which could not readily be resolved by conventional techniques, and elevated serum tumor markers, 15.

Now once again, we have at Memorial adopted the rather liberal policy to imaging patients with tumors and have considered whether -- have not used reimbursement as a criterion for whether we will do the patients, feeling that otherwise if we do that, we will impose a two-tiered system of health care on our patients.

I'm going to skip the first page of results, I think it's self evident what it is, negative and positive PETs by stage, and I want to

00127

go to the influence of PET on patient management. Basically, the point of this chart is that PET was used in the decision process to guide therapy in three-fourths of the patients. PET was ignored in 22.6 percent of the patients, and PET confirmed other studies in 3 percent.

I think this also reflects the type of patients that were referred. Over this period, you have to understand that at least 2,000 lung cancer patients, probably a similar number of colorectal patients were studied, so that the breast cancer patients that we see here are relatively small in number and these were the problem patients for whom conventional techniques were not able to resolve a particular management question.

Now, also, the next chart shows that the, whether the PET was negative or positive, did significantly influence, as you would expect, the actual choice of treatment or just watchful waiting.

The mode of therapy after the PET scan is in the next chart and I just want to spend a couple minutes with this. And again, Mr. Chairman, I did, there are some changes I

00128

would like to give you in terms of for accuracy's sake, because I notice in the first written statement, there is a summary which, for which some of the numbers were somehow miscopied, so I

will provide these numbers to you.

Basically, what this shows is the six-month condition, which again was our gold standard of whether the patient was stable or worsening, versus a negative PET and a positive PET. And it's possible using this information as the gold standard to actually compute an accuracy rate in terms of how the information influenced appropriately or not appropriately the choice of management. Now, we do this by considering that the negative PET with a stable treatment was essentially a false negative, because in that case there was evidence at the time of the therapy that there was disease and so the PET was disregarded, if you will.

So we're using essentially the clinician's judgment, putting all together the information, as the kind of gold standard for this particular study. So, there were 19 patients in that category with the negative PET, stable treatment, and we call that false negative. The

00129

negative PET with worsening in treatment, we call that a false negative, there was only one there. The negative PET who was treated conservatively but was worsening, that was a false negative, there were 10 there. The negative PET with a conservative but stable, a true negative, there was 28 there. A positive PET with stable treatment, a true positive, there were 38 there. A positive PET with worsening treatment, a true positive, there were 26 there. Positive treatment with conservative management who was worsening, was considered a true positive, and there were 4 there. The positive PET with conservative who was stable was considered a false positive, there were 7 there.

So this is quite a conservative way to look at the accuracy of PET. But the bottom line essentially from our study was that if we use the six-month follow-up as an indication of gold standard, the accuracy of PET for a guide to

management was 78 percent. Now I think that this should be compared with the fact that the conventional techniques were largely equivocal, that we know we will miss with PET significant microscopic disease. But I submit to you that in

00130

this group of patient, with all its limitations of a retrospective study, that data does support the view that PET can be useful in the management of patients.

So, on behalf of ASCO, I would like to thank you for your attention.

DR. PAPATHEOFANIS: Thank you, Dr. Larson. Any questions from the panel?

DR. BURKEN: I have a question. When you go back to the page here on characteristics of study patients, there were four indications listed for the PET scan, the conventional studies were equivocal, 63 patients for staging restaging, 33 patients, and so forth. I'm wondering, the data that's in your table and the following results tables are aggregate data cutting across all the indications; is that correct?

DR. LARSON: Right.

DR. BURKEN: Okay. So my question is, you know, are we doing ourselves a disservice by having these four indications lumped together in the results table?

DR. LARSON: I think it does lump together significantly diverse groups of clinical management questions. But we did it as a summary

00131

style and to give a flavor for the types of indications that were used on the request for patient imaging studies that came from the clinicians, so that we would do these studies.

DR. BURKEN: Thank you.

DR. PAPATHEOFANIS: Barbara.

DR. MCNEIL: Steve, I'm having a little bit of a hard time following this table, so where you said the accuracy was 78 percent, whatever the

table number is, can you just tell me what the associated sensitivities and specificities were for this table titled mode of therapy after the PET scan, do you have that handy?

DR. LARSON: The sensitivity that was calculated was 84 percent.

DR. MCNEIL: And the specificity?

DR. LARSON: The specificity, I'm sorry, Barbara, I don't have that number immediately available, but we can go over this later, and again, I will provide the correct numbers in the face page, because I noticed that there were some errors in the tables.

DR. PAPATHEOFANIS: Great. Thank you, Dr. Larson. I was looking at the agenda. We were scheduled for lunch from 11:30 to 12:30, we are

00132

now 15 minutes behind. I know the open public comment section is coming up at 12:30 and there are folks who have traveled from a great distance and I don't want to exclude any of that period of public comment, and so in that spirit, let's plan on a 45-minute lunch and let's meet and resume this session at 12:30.

(Luncheon recess from 11:46 to 12:40.)

DR. PAPATHEOFANIS: Welcome back. Let's regroup and get started. We do have a limited period of time together and I don't want to waste any of that.

Just a couple of comments about what is coming this afternoon for the panel. If you look at your technology assessment report book, you will see a series of five questions for the MCAC DI panel on FDG-PET in breast cancer, and after you've heard some additional evidence and after we have all had a chance to discuss the evidence, we're going to ask you to vote here on those five items, which will capture the four applications that were outlined in the technology assessment, specifically the applications being diagnosis, staging of axillary lymph nodes, recurrent and distant metastasis, and response to treatment.

00133

There will be several opportunities for the panel members clearly to discuss here their vital concerns. Also, with most of the speakers that you have already heard still here in the audience, there will be an opportunity for us to ask questions of those presenters who've already stood up and spoken. So, just to review, our goal and our charge is to provide HCFA with a set of recommendations or advice, as the title of this committee includes the word advisory, and we will frame that advice according to those five questions.

We are not going to make a coverage decision, we're not going to put policy into motion, that is the role of HCFA. We are a group of experts in our specific areas of expertise or specialty, and we are basically going to look at the available evidence, review it, discuss it, and then offer a specific recommendation.

So, at this time, we're going to move on to the open public comments section.

MS. ANDERSON: At this time we're going to open the mikes to the open public comments. I do remind any speakers who do assemble at the microphones to please state your name and your

00134

financial involvement with manufacturers of any products being discussed or with their competitors. You will have approximately three minutes in which to deliver information to the panel. We can begin.

DR. CONTI: I'm going to read a statement that you all have in front of you, or should have in front of you, from the Society of Nuclear Medicine, American College of Radiology.

Good afternoon, Mr. Chairman, members of the advisory committee, and ladies and gentlemen of the community. My name is Peter Conti, I am an associate professor of radiology and clinical pharmacy and biomedical engineering

at the University of Southern California. I am currently the director of the PET center and of radiology research at USC and have had over 20 years of experience in PET studies on cancer patients, spanning three institutions including Memorial Sloan-Kettering Cancer Center, the Johns Hopkins Medical Institutions, and for the last decade at USC. I come before the members of this committee representing the Society of Nuclear Medicine and the American College of Radiology, organizations of which I have been a member for

00135

many years.

The Society of Nuclear Medicine represents over 12,000 professionals dedicated to providing high quality diagnostic and therapeutic services. Likewise, the American College of Radiology represents over 30,000 practicing radiologists and nuclear medicine physicians with the same goal. For over a decade, breast cancer patients throughout the world have had access, albeit limited, to whole body positron emission tomography. Some of these patients have had the benefit of having their imaging studies covered under private sector health plans while others have had to pay out of pocket for such studies.

Thousands of breast cancer patients have been evaluated with PET, but thousands more have been denied coverage. That has been incorporated into the diagnostic practice in the cancer patient population, including those patients with breast cancer, in many facilities in the U.S. and abroad. Patients are referred by medical oncologists and surgeons for indications such as primary lesion detection, axillary staging, metastatic work-up, restaging and assessment of therapeutic response.

00136

As of May 2001, there were over 2,500 breast cancer patients reported in the literature who had received PET scans for diagnosis, staging,

treatment and planning, restaging, identification of recurrent disease, or assessment of therapeutic response. As the scientific program chairman of the upcoming Society of Nuclear Medicine's annual meeting, I can report to you that the data to be presented at that meeting increases the number of cases published in the literature to a total of 15 percent.

New studies to be presented in Toronto next week focus on staging and treatment planning, assessment and prognosis, measurement of treatment response, determination of tumor recurrence with restaging of disease, and those abstracts are attached to this document. These studies corroborate much of what has already been shown in the literature regarding the utility of PET scanning in this patient population.

The SNM and ACR recognize that much literature supporting the role of PET scanning in the breast cancer population may be technically limited as already discussed. However, no literature is without flaws or limitations. It

00137

would be inappropriate if not impossible to study every possible aspect or permutation of a disease or patient population prior to approving use of a new drug or medical technology for use in clinical practice. Neither patients nor their attending physicians would tolerate such a process.

On the other hand, patients and their physicians should expect a reasonable scrutiny and review of such advances prior to their acceptance into clinical practice. The challenges for regulators and providers is to identify appropriate indications and the threshold required for their acceptance.

PET is a safe procedure. The radio tracer FDG has been approved by the FDA as safe and effective for use in imaging cancer, including patients with breast cancer. It is shown to be highly sensitive, specific and accurate in the detection of many types of cancer as summarized

today. Of the breast indications noted above, however, the published peer reviewed data to support the use of PET in evaluating for residual and/or metastatic disease recurrence have emerged as the strongest clinically to date, despite what you heard in the first presentation this morning.

00138

We call your attention to three key full article publications from the literature.

A study by Bender of 75 patients looking at recurrence showed a sensitivity of 97 percent, specificity of 91 percent, and an overall accuracy of 93 percent. Notably, the positive predictive value of PET was 88 percent. And as an aside, I would say that I'm not sure I read the same article as was described this morning.

Another study by Moon et al. in 57 patients showed positive and negative predictive values of 82 and 92 percent in identifying recurrent or metastatic disease.

A third study by Huebner in 57 patients, showed a sensitivity of 85 percent, specificity of 73 percent, in the detection of recurrent or metastatic disease with PET compared to CT, where the numbers were 71 and 54 percent respectively, and mammography where the numbers were 2 percent and 100 percent.

Therefore, the recommendation of the Society of Nuclear Medicine and the American College of radiology to this advisory committee is to approve the use of PET at the discretion of the referring physician in the diagnosis of known or

00139

suspected recurrent or metastatic disease for the purpose of restaging patients with breast cancer. In this regard, we encourage the advisory committee to recommend that CMS consider the use of PET in patients who present with advanced breast cancer, when initial staging studies are required as part of the patient work-up.

The SNM and the ACR are grateful for

your careful attention to the needs of this underserved patient population and encourage you to adopt their recommendations so that more patients can benefit from this technology. Thank you.

I would also add as a personal note, the issue on the gold standard. This has been discussed at length earlier this morning, but I want to remind the advisory committee that the use or clinical follow-up is pervasive in the imaging literature as a method for assessing whether or not there is presence or absence of metastatic disease, and this has been extensively used in the PET literature as well, and should be considered as part of this evaluation. Thank you.

DR. PAPATHEOFANIS: Thank you, Dr. Conti.

00140

DR. WAHL: Hi, I'm Richard Wahl, I'm a professor of radiology at Johns Hopkins and director of nuclear medicine, vice chairman of radiology there. I am conflicted in that I received honoraria from Siemens, ADAC and GE at different times in the past relating to lectures on PET, and through the acquisition of PET Net Pharmaceuticals where I was a consultant, I have ended with some kind of class Q, some kind of shares of CTI, and also I'm a medical advisor to Mobile PET Services.

However, I have had an interest in PET for some time. In 1989 I think I was involved in the first studies imaging breast cancer with PET, showing feasibility of imaging primary, locoregionally, metastatic and systemic metastases, albeit in fairly large tumors at that time, and saw at that time that particularly in soft tissue disease, PET appeared to be uniquely capable of defining lesions.

I'm also principal investigator of a study I wanted to mention to you, one that Dr. McNeil actually helped design, sponsored by the NCI, in which we're evaluating PET

prospectively for the staging of breast cancer to

00141

the axilla. I just wanted you to know, this study is not yet completed, however, we have completed accrual of patients and we accrued 360 patients who have gone on to validation of PET scan results by axillary dissection, we're in the data analysis phase and hope to have this complete within the next few months. So we hope that this will be the largest prospective study of PET in breast cancer staging, specifically for axillary disease.

We also are examining the prognostic value of PET in this group of patients by following them. Because of the gold standard issue and the variability of sampling of axilla, I think the tendency now is to sample more extensively small axillary nodes repeatedly and do staining which may upstage patients from a stage they were previously, so we think the prognostic part of this study is also very important. Anyway, I just wanted you to know that it is coming, but I don't have results.

I wanted to comment that based on my experience at the University of Michigan and now at Hopkins, I believe PET does have a definite role in breast cancer and indeed, I participated in a panel last Monday in Vancouver, British

00142

Columbia where the British Columbia Cancer Care Agency was trying to decide how do they use the limited resources in British Columbia and the limited access to PET in imaging breast cancer. Clearly they are resource constrained and are trying to rationally apply imaging methods. And I was asked to summarize the expert panel meeting with a lecture entitled, in what situations should we no longer be practicing oncology without PET.

And in the situation of breast cancer, this conservative assessment was that in particular, recurrent breast cancer assessment, particularly for soft tissue metastases, was a

unique situation that should be supported by the British Columbian government, specifically the situation of brachial plexus recurrence versus radiation necrosis, which is a very difficult diagnosis to make, and also the chemotherapy response assessment in patients with large primary breast cancers and in follow-up known breast cancer were viewed as indications where the literature was sufficient to support the implementation of PET. Other areas were felt in need of further study.

I did want to comment particularly

00143

about Mr. Samson's comments. He did discuss the study that I did in 1993, reported in 1993, about PET in following treatment response. He indicated that there was a question as to whether the patients, whether the persons reading the mammograms were blinded. Indeed, they were. The PET scans and mammograms were not used for management of the patients and patients were managed by conventional methods because PET was a new technology at this time.

So in summary, I believe there is abundant evidence in soft tissue disease to support the use of PET. And for recurrence, I think one of the problems we face is that some of these conditions are very infrequent, the brachial plexus issue as an example, in about eight years at Michigan, we only had 15 cases, PET consistently performed more accurately than MR. We have a paper in press in the JCO showing this, and to get to a hundred patients is going to take many more years. At the time I left Michigan, it was impossible to get a referring oncologist to order anything but a PET scan in this clinical situation, so I would encourage you to look very carefully, and support the ACR SNM position, and

00144

possibly also very strongly consider the chemotherapy response data, which in over a

hundred patients is very strong. Thank you very much.

DR. PAPATHEOFANIS: Thank you, Dr. Wahl.

MS. PIERCE: Good afternoon, and thank you for the opportunity to address the committee. My name is Kim Pierce. I'm a breast cancer survivor and a member of the National Breast Cancer Coalition, the Coleman Foundation, and I am here in representation of the thousands of women who are diagnosed with this devastating disease annually. We received over a thousand signatures in two hours at the Race for the Cure for the Coleman Foundation.

Like many other women, I had the normal concerns about breast cancer, so I got my annual mammograms and physical examinations and I performed self exams in between, and like lots of other women, when I discovered a lump in my breast, I had the standard tests performed all over again, as well as ultrasound, but when the results came back negative and my doctor told me that we would just wait and watch, I felt

00145

relieved. After two years of negative mammograms and ultrasounds, I became increasingly concerned about the lump because it was continuing to grow. That's when I heard about PET imaging. Fortunately, I worked in a medical center that had, and I had access to PET.

When the PET scan showed that the other tests had been wrong and I did have a malignant tumor in my left breast, I was immediately scheduled for biopsy which confirmed the malignancy was infiltrating lobular cancer. Infiltrating lobular cancer is not routinely picked up by mammography, but because PET revealed my tumor when nothing else did, I was able to get the treatment I needed in time.

Unfortunately, until HCFA approves PET for special cases like mine, where mammography and other tests are not effective, more women will

find out that they have breast cancer too late to be cured. Most women have never heard about PET, because it's not available to them for diagnosis or staging of breast cancer, even though it is one of the most accurate tests available to women with dense fibrous breasts, women who have had medical or cosmetic surgeries, or even biopsies performed,

00146

or women like me with a form of breast cancer that mammography cannot detect.

There are many other women, with the numbers increasing each year who have had their breasts scarred by various procedures. This causes problems for mammography and palpation. While all of these factors alter the accuracy of mammography, CT and physical exam, they do not interfere with PET. Its high energy radiation easily passes through these tissues so that PET can differentiate benign processes from malignant ones. I believe that PET is extremely valuable in diagnosing women in those subpopulations for whom other screening technologies are less effective.

PET can also appropriately stage breast cancer patients by showing axillary and mammary nodal involvement and/or distant metastasis in other organ systems such as bone, liver, lung and brain, all in a single examination. This can change the treatment of breast cancer and spell hope to more women with their terrible disease.

I have met hundreds of women who were inaccurately staged at diagnosis and therefore, did not get appropriate treatment. These women subsequently died of breast cancer. I sincerely

00147

hope that MCAC and HCFA will understand the benefit of PET for women like me, so that potentially life saving and cost effective medical technologies are made available to the female Medicare beneficiaries who need them. Thank you.

DR. PAPATHEOFANIS: Thank you, Miss Pierce.

DR. WEINBERG: Hi. My name is Irv Weinberg. I'm a radiologist and physicist. I was trained in oncology imaging at Johns Hopkins Hospital, built the first dedicated device for breast PET at the NIH, subsequently took the entrepreneurial route in developing dedicated instrumentation for PET breast, and I am now president of PEM Technologies.

I would like to highlight the possible effect of your decisions and your language on emerging technologies. We are focusing on methods of diagnosing extent of breast disease. The technology itself has been published in the European Journal of Nuclear Medicine, Journal of Nuclear Medicine, Medical Physics, it is very clear from the point of view of physics as to possible advantages of this emerging technology.

If there is any cancer that requires

00148

physiologic and biochemical imaging, it's breast. This is an endocrine disease, it is exacerbated by reproductive histories that affect endocrine status of the patient. It is treated and prevented by hormonal therapy, it is clearly an endocrine disease and requires biochemical imaging.

I would just appreciate your sensitivity to the future or emerging technologies that may represent the application of physiologic and biochemical imaging to breast disease. Thank you very much.

DR. PAPATHEOFANIS: Thank you, Dr. Weinberg.

DR. ALAVI (phonetic): I am Bahs Alavi (phonetic), I am professor of radiology and chief of nuclear medicine at the University of Pennsylvania, and I work with ADAC as a consultant to them, and my group also deals with them for instrumentation.

The idea of FDG came about in 1973 at Penn, and in 1976 we administered the first dose of FDG to human beings. So 25 years later, we're

still arguing about the role of FDG, while MR was around for no more than two or three years and was

00149

approved for funding. So it's nice to see that there is a discussion about applications of FDG which of course for someone like me who has been with it since the beginning, I am actually happy to see the data that FDG has come along so far.

I do of course a lot of patients every day, 10 to 12, and a variety of disorders, and I truly believe that the role of FDG in cancer has been revolutionary. In particular, I would like to just mention a study that I was funded by the Army to do in metastatic breast cancer who were candidates for bone marrow transplant. There was the (inaudible) study to see whether we can predict who is going to respond and who will not, since only 20 percent of the patients will be cured by bone marrow transplantation.

A side finding of the study was to compare FDG with other imaging modalities, which included everything that we do for cancer, namely chest x-ray, bone scan, CT scan, as part of the study. We enrolled 39 patients and most patients had more than one study, so we had to analyze our data, and our results indicate that one FDG stand-alone could be equal to all the diagnostic studies except that bone scans appeared to be a

00150

little more sensitive than FDG.

(Inaudible) flaw of the bone scan, because we usually see longstanding effect from cancer in the bone, it lasts for a long time, and that really gives us an indication that disease is active, that FDG shows some of those patients not having active disease.

So I believe that this is going to be an effective technique, especially with metastatic cancer, doing one single study allows you to look at the entire body in three dimensional space, versus doing a CT scan for the liver or bone scan

for the bone, so if the other diseases are an indication, which I think they are, FDG is going to be the study of choice for metastatic disease. Thank you.

DR. PAPATHEOFANIS: Thank you, Dr. Alavi. Anyone else?

Anyone on the panel that would like to recall any of the speakers for any questions at this point? Okay.

Anyone else then who spoke before the open public session that may also want to address the panel at this point? Okay. If not, we're going to move on to an open panel deliberation,

00151

and I think the best way to start is to quite literally go down the list of five questions that HCFA wants us to address, and so why don't we spend our discussion along those lines and let's start off with the first question, is there adequate evidence that PET can improve health outcome when used to decide whether to perform a biopsy in patients with an abnormal mammogram or palpable mass? Jeff.

DR. LERNER: Frank, I have a question actually before we go directly into going through the questions. One of the things that I guess surprises me a little is in the prepared presentations and in the open public comments, there wasn't to my mind a great deal of critique of the TEC assessment, and at the same time what I think a lot of the public comments had in common was that they were more looking at in a sense Medicare policy, you know, how we make decisions, as opposed to looking at what I interpret to be the direct charge of this committee which is to go through those questions. And you know, I'm not quite sure what to do about that, but I think it's important to raise that issue because I don't want to seem unresponsive to what the audience has

00152

raised, because if we just go through these

questions, at least to my mind so far, I feel these are fairly clear-cut.

So I would like to at least ask the question, whether people have questions about the fundamental assumptions going on, and there's some follow-up to that, but why don't I leave that for the moment.

DR. PAPATHEOFANIS: Sure. Sean, can you speak on the process by which this technology assessment came to be, and sort of the internal events? Maybe that will get us started.

DR. TUNIS: Sure. Actually the technology assessment, this particular technology assessment was already in process before the HCFA had decided to refer this issue to the MCAC. It was being done and Carole or Debbie can correct me if I'm wrong, was being done for the purposes of the Blue Cross/Blue Shield Association medical advisory panel to make their own recommendations about coverage in Blue Cross/Blue Shield. That was the reason this TEC assessment had been started.

As part of our review of the coverage request from July of 2000 for broad coverage of

00153

PET and when we concluded that review in December of 2000, had ended up extending coverage for four additional cancers, I believe it was, to a total of six, and at that point had been decided that several issues would be referred to the coverage advisory committee, this issue being one, and then we worked with the AHRQ to piggyback on to the work already being done by Blue Cross/Blue Shield to have this TEC assessment ready in time for this meeting, so that was this process.

I don't know Frank, or Jeff, if you wanted me to comment more broadly on sort of the role of the MCAC in this process in terms of the focus on the evidence versus the sort of policy and the thresholds for decision making.

DR. LERNER: Maybe I can help a little bit by just making one more statement. I almost

found the public comments, that they would have been more useful in the entire coverage process if they had gone in prior to the formulation of questions, and maybe they did, maybe other things went in there, but by the time we reach this stage, as I understand the charge of the panel, is to answer these questions, and we certainly are prepared to do that. But I think, what I'm

00154

wondering is whether the audience and the people who commented will feel that that is responsive that they've been heard, because they raised all kinds of issues. I have my own list and I'm sure other people do.

DR. TUNIS: Well, maybe a question to ask that would be a clarifying question, again, anyone from the public who has spoken can address this, is, I had gotten the sense that at least several of the speakers were not contesting the fundamental conclusions of the technology assessment, which for the five questions asked here were negative conclusions in terms of adequacy of evidence. So maybe, I'm not proposing that that's a correct restatement of what folks have concluded, but maybe if there are folks who have spoken who believe that any of the conclusions in the TEC assessment are in fact incorrect, then maybe we can get that conversation moving further by addressing that explicitly.

DR. PAPATHEOFANIS: Sam, before you go on, is this what you're getting at, Jeff?

DR. LERNER: Yes, it is.

DR. PAPATHEOFANIS: Okay. Go ahead, Dr. Gambhir.

00155

DR. GAMBHIR: You know, first of all, I think the TEC report is done in a very professional manner, very rigorous in its design and its actual reporting of results. I think the problems I have with it as well as other people are when you're looking at a new technology such a

as PET in the role of breast cancer, the first question is should the inclusion criteria for studies be what this particular report chose as the inclusion criteria?

For example, as a lot of people have argued throughout the day, there are other articles that don't meet the inclusion criteria but exist in the literature. For rapidly emerging technologies, just like we've argued in the past, there continue to be abstracts that emerge which will eventually see publication but have not seen publication yet. To us, those need to be weighed into any emerging technology report, because it strengthens the confidence for the N in the case, number of patients or number of studies performed. So the one area when I read the report and was actually asked to critique it, the one thing I thought would be useful is to actually include abstract.

00156

That's why in my presentation what I tried to show you is that when you start to include abstracts, and of course you can't use the inclusion criteria then, because one of the inclusion criteria is it be a research article, but when you start to use abstracts, the sensitivities and specificities all remain in these same ranges, but the confidence goes up, because now the number of patients, as you saw in most applications, is almost doubled. And that doesn't even include abstracts that have just started to come out or are due out next week.

So I think one problem we have with the report is how to be fair to all the literature and how to be fair to abstracts specifically.

The second problem for the report is that although we agree that the conclusion if you only include those articles show there is limited evidence, if you start to include the other papers and abstracts I'm talking about, we think there is strengthened evidence for these other applications. If we focus on recurrence and we

focus on staging after recurrence, or monitoring for therapy, the numbers almost double from the numbers presented previously. To me, that adds

00157

confidence in those accuracy values. And as I stated, I don't think the issue is what is the sensitivity and specificity of PET for this particular application.

You can revisit these over and over and over, and just like we did with lung cancer, you will see them converge into a range with increasing N, and they stay in that range. The bigger issue is, given those accuracies and the clinical management algorithm, how many good benefit outcomes will you have for your patients and how many harmful benefits. And that's why what I tried to show was that if you look at certain underserved women that are not served well in the current management algorithms, we think PET's useful.

So I think those are the issues, but I don't have a problem of saying if those are the inclusion criteria, although we might disagree with the gold standard issue and by the way, did the other result know about the PET results, did the other biopsy know about PET, that's sort of a misunderstanding of what happens clinically, that pathology reports don't need to understand that.

DR. MANYAK: You know, I have been

00158

struck today with something that I was unaware of reviewing this literature regarding breast cancer. Since I don't deal with breast cancer very much, being a urologist, we avoid it, but we have certainly some parallels in our field as well with the diagnostic dilemmas that are faced here. And the thing that struck me here today is that there is a subset of patients where the question hasn't been asked, and it's not because of the fault of the construct of the technology assessment group, but it's a question that I'm not sure, I don't

know if the other panel members were aware of it, certainly one I wasn't aware of, and that is that there may be a subset of patients where this does show a greater benefit than what's existing out there, such as your dense breast tissue patients.

Now that raises questions in my mind, what defines a dense breast, and avoiding any jokes or anything else, seriously, is there some measurement of that and first of all, is that universally accepted and is it universally applied in clinical settings, and if it is, what's to prevent the use of PET scans to escape outside a dense breast tissue patient.

I mean, these are all issues that come

00159

into play, but if you pick out a subset where there may really be an advantage to PET, and it may be with that subset, I have heard several people mention that today, but that data we couldn't glean from the literature, and I don't know if it exists in the literature. Those of you that really looked at this very carefully may be able to answer that.

DR. GAMBHIR: Yeah, let me clarify that. So first of all, of all the applications we've heard, there is the screening category and then of course the management after diagnosis. In the screening category, first of all, dense breasts is an artifact of mammography, that is, if you had a world where for some reason mammography never existed and PET existed before mammography did, we wouldn't be talking about from a PET perspective dense breasted women and non-dense breasted women, because as I said, PET radiation doesn't care about density of breast tissue.

There is a formal way to grade breast density. It is published in the literature and is called the Wolf grade. There are four grades of breast density, with DY, the category I chose in that decision model, being the densest of the

00160

dense breast categories. Grade DY women, of which there are estimated to be about 3 million on the high end, and on the low end 500,000 women, are the kinds of women that as I argued, are underserved by mammography. It's now no longer a question of oh, how many biopsies did you avoid or is there a harm from not catching something.

Those women are being harmed now because in fact, they are screen, nothing is detecting anything, they go back, have their next screen, their next screen. In the decision models I would love to be able to show you that oh, there's a trial comparing only dense breasted women, mammography versus PET. It was asked to me outside, why hasn't such a trial been done? Part of the reason is because it's such a low incident of breast cancer in the screening population, to do such a trial takes a long time to pick up dense breasted positive findings. So it would take years, literally five to seven years to get even enough N in those women.

But the second thing is, remember in my reasoning, the dense breast stuff is an artifact of mammography. From the PET world, there is no difference in response for the signal from dense

00161

breasts versus normal breasts. We have just the same chance of detecting a lesion within a dense breast or normal breast. Where is that evidence? That evidence is in all the literature we do, the normal and dense breast women are both screened, all the data you see presented, it's not like we're subdividing it into dense breasted versus non-dense breasted women.

So I think that's one area where although no clinical trial exists, it's proven head to head that if you take a look at a decision model, use good judgment based on what data is available, there's likely to be a useful benefit for that subgroup of women.

DR. LERNER: Can I ask you a question on that, Sam? When we talk about the Medicare

population, we should be clear about who that is, but if you said it is people over 65, how frequent is the dense breast issue in that age group?

DR. GAMBHIR: Certainly it's much higher in the younger age group than it is in the older, but I think we shouldn't think about it in terms of well, will this affect the over 65 population from a reimbursement point of view, because what's done here is of course watched by

00162

all kinds of providers. So I think the issue is, dense breast women of any age are being underserved, and if you say which dense, where are more dense breasted women, younger or older, it's more younger women that have dense breasts.

DR. LERNER: But you see, it does go to the charge of the committee, and for purposes of this being a Medicare committee as opposed to you know, a committee for the whole population of the United States.

DR. GAMBHIR: Right, but I'm saying what is done here is watched by more than just --

DR. LERNER: Yes, I agree.

DR. GAMBHIR: So if we say what women are being underserved, it's women of all ages with dense breasts. The fact that there's less women that are older with dense breasts is a relevant issue to some of the direct reimbursement from Medicare, but it's not the only issue when we look at which women are underserved in the entire population, which includes all dense breasted women.

DR. PAPATHEOFANIS: Sure. Donna?

MS. NOVAK: It sounds like we're really talking to question five here, is that correct,

00163

that question one and two assume that there has already been a mammography, and question five is saying, is PET an alternative to, a better diagnostic, am I interpreting that question correctly?

DR. PAPATHEOFANIS: No. This is a whole separate issue really, and I think we have gone probably a little farther than we want on the dense breast issue at this point.

MS. NOVAK: Okay. Well, I guess my question is, if we are first to assume that there has been a mammography.

DR. PAPATHEOFANIS: Yes.

MS. NOVAK: Five does not, if I read it correctly, and I guess where my question was going is, one of the things that surprised me is that we didn't see any evidence at least that stuck out to me as to, you know, if PET is really a better diagnostic tool than mammography, which we kind of always assumed that it has been. Is that true?

DR. PAPATHEOFANIS: Is that what you want to speak to, Barbara?

DR. MCNEIL: Well, yeah. I have just a procedural question and maybe it's to Sean or to Frank. I'm getting a little confused about what

00164

our charge is and what we're supposed to do, because I want to make sure we do the right thing here and we use the right information to make it. So I read our little bible here about recommendations for evaluating effectiveness this morning again, and this tells us that we're supposed to give you Sean, and HCFA, advice about the evidence.

So my problem is, there is now an indication that's on the table for which we have no evidence, and I am not sure that given this statement, that I personally feel comfortable about making a judgment in the absence of somebody giving me some data other than comments. And part of the reason I got more worried about this than what I was this morning, because I could see that was coming up on the agenda, is the fact that I guess Steve or somebody raised the issue about the potential for biopsies, false positive biopsies, unnecessary biopsies impacting subsequent mammograms, somebody over there.

So that made me think, well, we can't assume that every positive PET study is a true positive, I don't think, because we know we have some specificities that are not 100 percent in all

00165

of these indications. So if that's the case, then we know, or it would be reasonable to assume that there would be some false positives in dense breasts, just following the same line of reasoning.

And then taking up on the question that I never would have thought to ask this morning, in a million years I wouldn't have thought to ask this about biopsies, then I'm now wondering about the impact of those on this whole discussion that Sam is raising. So this whole, what I am trying to say is, I'm feeling very uncomfortable personally about getting into any of the data on this subject, because we have no data, and I would almost propose that this is a question that we can't answer today.

DR. GAMBHIR: Let me just answer that by saying first of all, there is data. I think we're getting confused about the data that's out there. There is data on FDG-PET in detection of the primary breast tumor, both in screening studies as well as in palpable masses, as well as nonpalpable and palpable. So it's not fair to say --

DR. MCNEIL: But it is not here, Sam.

00166

It hasn't been presented to us.

DR. GAMBHIR: Actually, no. Even in the blue TEC report, when you look under the diagnosis category when they're talking about looking at the primary and lymph node staging, the primary detection data is the data we're talking about. That data is there.

DR. FLAMM: Except, I think there is a clinical difference when a physician refers for a focal abnormality and a focal evaluation, and

someone coming in off the street for a screening study.

DR. GAMBHIR: There is, but the abstract data and other data which, you know, may not be fully in the blue TEC report, but the other articles I showed do in fact show even those populations, that is, people walking off the street, the screening groups, so I don't -- I wouldn't say that this is out of the blue that you know, there is no data on this, or we just said let's pick on dense breasted women. The reasoning is, to try to find an underserved group that would benefit, say what is the existing data that's applicable to that group, and what I'm trying to argue is that from the PET perspective, all these

00167

women that have been scanned where we were looking at the primary lesion, it doesn't matter whether they were dense breasted or not, so that data applies to that decision model, and that is the key issue that links that data to the model that's in breast cancer research and treatment that was originally designed to answer this question, what is the role of a second study inserted in when a first study like mammography does so poorly.

Now I realize from your perspective it's frustrating to say, but that isn't one category that was addressed specifically in the report, but I think it's a category we need to visit, because it's one of the most important categories from a perspective of women that are currently underserved.

DR. TUNIS: Let me just address and try to at least clarify from my view procedurally what we should try to do taking this into account, and you know, I think this is going to stay a little bit confused, in part because there is an important new issue that's been added to the table and we have to figure out what to do about it, and that's the dense breast issue. The charge to this committee is in fact to review the evidence and

00168

essentially the framework that we are ultimately going to go through is to answer these five questions, around which you have the five questions to the panel. So we will do that and we will take a vote on those five questions.

The issue of you're supposed to consider the evidence, as Janet said at the beginning, we think about the evidence broadly, so the evidence is what you got in advance and then whatever else people bring into the room to your attention, including what Dr. Gambhir has raised and what other folks have raised. It's new evidence but it's still part of the evidence. It may not be published evidence but it's still a form of evidence and you still have to deal with it at some level, so we will deal with that issue. We won't take a formal vote on the issue of dense breasts, because it's not one of the questions that we were sort of in advance charged to answer, but we will continue to discuss it.

So, I don't know if that clarifies things but at least, we will go through an orderly vote on the issues on the table before us and I think we will be staying within the boundaries and the guidance of the panel in terms of considering

00169

the evidence.

DR. PAPATHEOFANIS: Before we go on, there are a couple of things on the table right now. The first question you had, Jeff, and the dense breast issue hass sort of now become the focus, your first question was a critique of the technology assessment. Is there anyone that wants to provide some discussion on that, and afterwards what we will do is return to the issue of the dense breast and as you just heard from Sean, we will not be voting on this, because it is not an issue that we've had a chance to really spend some time and have been provided any sort of background material on.

What I would want to do is open the

floor so that each of us can provide any comments regarding their personal position or opinions or thoughts on this dense breast issue, which may be revisited at a future MCAC panel meeting, but I want to just finish with the issue of critiquing the technology assessment, and Dr. Zarin?

DR. ZARIN: I just thought I would explain where the five questions came from, because what we're talking about now is really a sixth question or a subpart of one of the other

00170

questions, depending on how you look at it. The questions came from ongoing discussion between HCFA staff and the people who had applied for coverage, as well as other interested parties, between us, the Agency for Health Care Research and Quality, Blue Cross/Blue Shield TEC and HCFA staff, and they were really designed to reflect what we were hearing were the proposed indications for PET scanning. So they weren't sort of arbitrary and they weren't simply what Blue Cross/Blue Shield decided to look at, but were based on what we were hearing were the proposed indications.

And the specific questions came from applying the MCAC Executive Committee's criteria, the bible as Dr. McNeil referred to it, as to these indications. So that's where the questions came from. I think the issue of dense breast is raising the issue, as Sean said, of how to deal with sort of a new indication that comes up at the time of the discussion, and there wasn't a systemic assessment of that indication, but that's because it hadn't been raised ahead of time.

DR. PAPATHEOFANIS: So from your perspective as the chair of the technology

00171

assessment group for AHRQ, is this a typical product that you can expect from the EPCs and is it in keeping with those standards?

DR. ZARIN: Well, the question is, how

do you determine this sort of a policy question of what to do about coverage for PET for breast cancer, and that has to be kind of turned into a set of research questions, if you will, and that process is a very key process, and this was done collaboratively between HCFA staff who were in contact with the different stakeholders, as well as those of us who were reviewing the actual data. And we did it as best we could to try to come up with the indications that seemed to be being proposed and which seemed the most promising, sort of the best case argument for the use of PET in breast cancer.

I think what we're hearing today is given the findings there, as people's thinking has evolved perhaps, maybe one of those questions has been refined further, and maybe that's unavoidable. I'm not sure if that could have been known several months ago.

DR. PAPATHEOFANIS: Great, thank you. Anyone else that would like to comment on or

00172

critique the assessment? Dr. Phelps.

DR. PHELPS: I have a question about procedures actually, because I think the dense breast issue is a paradox, because to mammography and palpation there are dense breasts but to PET there are not, it's the diagnosis of breast cancer. So you know, I think with that paradox, the committee has to determine, has to rule about whether dense breasts fit into PET's criteria of diagnosing breast cancer or their radiographic palpation criteria that makes them a subpopulation, so I would ask you to do that.

DR. PAPATHEOFANIS: Great. Dr. Conti.

DR. CONTI: With all due respect to the comment that was made earlier, I'm a stakeholder as in the Society of Nuclear Medicine, as is the American College of Radiology. We were not consulted on the nature of these questions, so I beg to differ with that comment. I'm also not aware of any other stakeholders in the audience

from other professional societies here that were consulted on the structure of these questions, so I would like some clarification on that.

Now specifically with regard to these questions, I would also like clarification on what

00173

health outcomes means, because I think if you as the majority of people in this room how you would want to evaluate diagnostic imaging technologies, health outcomes would probably fall to the bottom of the list rather than the top. We're looking for management changes, we're looking for decisions that are made in respect to the introduction of the procedure.

Health outcomes are in large measure in breast cancer patients determined by the treatment choices that are made, and those made by the surgeon or the medical oncologist, so we also need to be clear what those measurements are. And I don't believe that the questions reflect the reality of diagnostic imaging measurements, and I don't think they reflect the technology assessment that was done, because that wasn't addressed at all as far as I can see.

And third, I would also point out that in my statement, we specifically presented arguments that go contrary to the results of the technology assessment with regard to recurrent disease and metastatic breast cancer, and it is documented for you.

DR. PAPATHEOFANIS: Thank you.

00174

Dr. Wahl?

DR. WAHL: I did not have an opportunity to review the Blue Cross TEC report much before this meeting. I did get a look at it. But I had an opportunity to review personally the breast cancer PET literature in writing a review article for the Seminars in Radiology, and this will be coming out shortly, so I did take a very careful look at the literature, including

abstracts, and I do believe one of the limitations of the TEC report is not looking at abstracts.

Further, specifically regarding questions four and five, my read of the literature and my conclusions in my review was that clearly, PET is in virtually every study in which it has been examined for looking at distant metastatic disease, it performs as well or better than conventional methods, and as a single test could replace several other tests. So the question was, could it replace standard imaging tests? It's hard for me to say if the accuracy is as good or better, that it couldn't.

Similarly, the fifth point, and I just wanted to comment that of course the difficulty in doing studies in metastatic disease is that you

00175

certainly can't biopsy every normal tissue, so it's very hard other than follow-up, to determine what is true in these studies. So the situation in determining assessment of accuracy of metastatic disease is really hard. So my conclusion in my review is that the fourth point, I would certainly differ in the conclusion, and I just wonder if the entry criteria in the TEC assessment are completely appropriate.

The other question, number five, I know that one of my studies was quoted, the one from 1993, which was the first to prospectively look at PET in assessing the response to chemotherapy. It was described as having two PET scans in each patient and in fact it had five PET scans in each patient, sequentially done at base line, day 8, day 21, 42 and 63, looking at the time course of change in PET compared to independently and blindly read mammograms. And what that study clearly showed, it was in the JCO in 1993, statistically significant was that PET showed a much more rapid change in response to effective therapy than did mammograms. Mammogram didn't change in this period of time, so conventional diagnostic methods didn't change, and the PET scan

00176

changed very rapidly and did significantly by eight days after treatment, with further reductions in metabolism with additional treatment.

So that, that wasn't discussed but that was one of the questions, it does provide an earlier response assessment than conventional response criteria, and that was specifically in a paper that I don't believe was accurately quoted in the review. Again, I didn't read the entire review, but at least in the summary presented today, and I think that's consistent with other studies.

The other concern I had about the review is as regards the fifth point was that there was an emphasis on denying patients therapy in case PET was falsely showing a lack of response. Indeed, PET showing response much earlier than mammogram or measurements of tumor size, I think that's improbable that it would happen, that it is a more sensitive measure of response.

The other concern not addressed was what if you treat a patient too much with aggressive treatment, some of those drug regimens

00177

contain six drugs, who aren't responding? I think it's a tremendous disservice to a patient. And not including that argument and not assessing the relative weight to that potential damage I think would be a limitation in the analysis. I wanted to mention that I did have those disagreements based on my review of the literature, and I would be happy to provide you with a copy or preprint of that Seminars article if you need it, that was recently completed. Thank you.

DR. PAPATHEOFANIS: What I don't want to do is have another session of open public comment. I would really like to hear the thoughts of committee members. Jeff, have you heard enough

as far as critique of the technology assessment at this point?

DR. LERNER: Yeah, I think so.

DR. BURKEN: I need to make a comment in response to Dr. Conti in terms of the formulation and design of the questions. The questions were really designed as a combination of CMS as we call ourselves now, the Center for Medicare and Medicaid Services, and I'll try to stick to that if I can, between CMS and AHRQ, okay.

00178

We have become increasingly through web site postings, but not everything we do is totally transparent, and Dr. Tunis may want to kind of respond in which directions we may be going or not going in terms of transparency. But as I said, it was not a fully transparent process, nor intended to be, for formulating the questions.

DR. PAPATHEOFANIS: Thank you, Dr. Burken.

Well, with the critique of the TEC assessment off the table at this point, and we can return to it if there is a need or if there is time, I would like to refocus on the other issue that snuck in on the table so to speak, and that's the issue of dense breasts, and I'd like to hear from the panel members. And again, I welcome you to ask for audience input, but I think we're beginning to get a flavor of what that input will be, and I'd rather have you share some of your thoughts as this is an opportunity for you to do so. Mike.

MR. KLEIN: Okay. One of my observations is that the issue we have been debating or at least has been on the floor here, is so much of how one defines what is the disease

00179

that we're dealing with. And some of the comments that have been made have been along the lines of looking at the, the need to look at breast

anomalies, in this case cancerous lesions, biologically or has been described as an endocrine problem. And as such, the imaging technology that exist today don't effectively, it would appear from discussions, don't appear to address the biological aspect of it as such.

Functional imaging or biologic or metabolic imaging is the issue, so I contend that the dense breast issue is a part of that. If you reduce it to just a dense breast issue, you will introduce the issue of ultrasound, which is certainly a viable way in conjunction with mammography of looking at and diagnosing dense breast tissues. Certainly in Asian countries where there is a very high incidence of dense breast tissues at all ages, ultrasound is not only used as an adjunct to mammography for dense breast tissues but is in fact in many areas used as the preferred method of imaging.

So I think it's part of this issue of looking at it as more of a biological disorder than one that needs to be treated as such, and I

00180

would be interested to make some additional comments later when we talk about how PET can be used in the staging of the disease, treating it, and certainly for recurring and for other risk factors. But I'm not sure that the dense breast issue in and of itself is the central point that was being made by the speakers. I think it was the issue of this is more of a systemic or biological problem. If someone wants to comment or correct me on that, please do so.

DR. PHELPS: I think if you just stop for a minute and look at the very signal, you know, where is the signal coming from in x-ray techniques, and even through palpation addresses it, and even ultrasound, those are all issues related to the density, so the very signal that you're collecting to make a diagnosis is coming from density. And when you turn to PET, it's not the fact that the imaging can penetrate that

tissue easily, which it can't, but the signal is not coming from density, it's coming from the glucose metabolism so it has nothing to do with density. Density happens just to be in the clinical work-up by both palpation and the x-ray techniques, it happens to subpopulate them, but

00181

they don't subpopulate in PET, because they are metabolically differentiated.

You know, that's why I was responding to Barbara's comment that it depends on how you want to take, the direction you want to take. If you say yes, I accept that argument, then they are not a subpopulation to us and the diagnostic criteria apply. If you subpopulate them by the density, then they are subpopulated that way and you might exclude them from the questions.

And I think you has asked the question actually in the beginning about you were concerned that some of the people were raising questions that were not in the questions here. Now I respectfully would say that this is a process in evolution so you know, there are mistakes that will be made and it's improving, and we also have to do a better job of engaging you, so next time we will do better on your side and our side about the questions, but there will be some mistakes.

DR. PAPATHEOFANIS: Dr. Weinberg?

DR. WEINBERG: Yes, if I may, just with regard to this dense breast issue and how it relates to biologic imaging. I think if you look at symptomammograpy, which I have some

00182

publications which I participated in some publications on, the question there is problem solving, and can functional imaging assist in problem solving in difficult mammograms. And dense breast is really one subset of difficult mammograms. It may be a patient who has had a biopsy in the past, it may be an elderly patient who is on hormone replacement therapy who all of a

sudden has a density that wasn't seen on the previous examination.

So I think the question of not only whether to perform a biopsy but more importantly for us is where to perform a biopsy on a patient with difficult mammograms is a very critical issue to every mammographer.

DR. PAPATHEOFANIS: Thank you. David.

MR. SAMSON: I would like to pose a question to the committee having to do with the breast density issue. In the technology assessment report, we tried to distinguish between two segments of the biopsy population, the upper segment that has clearly abnormal mammograms and palpable masses, and the lower segment that might have an indeterminate mammogram. And I wonder if there is a relationship between the lower segment

00183

and patients with dense breasts, whether there are patients who have a dense breast and have an existing tumor that is fairly large in size, would that be picked up in spite of the density of the breast?

And is there a lot of overlap between the, I guess the smaller tumors, the nonpalpable ones, indeterminate mammograms and the patients who have dense breasts? Are, the ones with dense breasts tend to be smaller tumors. Is that the same issue?

DR. PAPATHEOFANIS: What do you think?

MR. SAMSON: I don't know, that's why I'm posing it to the committee. And if so, if it is the same issue, if the dense breasts are hiding small tumors, then we need to know the diagnostic performance of PET for small tumors, and we don't know that. That's my point.

DR. FLAMM: I think there are some logical similarities. You have to think about patients presenting for mammography as being a whole spectrum of different types of patients, and we have diagnostic performance data in a very specific segment of that population. And I have

concerns about generalizing these diagnostic

00184

performance characteristics across the whole range of patients who present themselves for a PET scan. And I think we need to be very clear about what we know and what we don't know about the diagnostic performance. We can't just say these are diagnostic patients so therefore, we can take these estimates, because I think the types of lesions you would want to pick up in a patient presenting with dense breast de novo for her screening study would be different than someone who is coming in with a palpable mass for the PET study, to diagnose it as benign or malignant.

DR. GAMBHIR: Let me just respond to that. There is some --

DR. PAPATHEOFANIS: Sam, you have two seconds.

DR. GAMBHIR: In fact, when you have larger lesions, those can also be missed in dense breasted women. For example, it's not just the issue of lesion size and sensitivity in both mammography and PET relate to versus the density of the breast versus nondense. So the literature shows that in dense breasted women, even lesions that are larger in size -- the example I showed you was a one centimeter lesion that was missed by

00185

mammography entirely, actually on three subsequent uses. So it's not simply that oh, PET is catching those larger lesions and is going to miss all the small ones and that's really what mammography is missing on dense breasts. It's not that clearcut. Now there is an issue of exactly what is the sensitivity and specificity of mammography, PET, ultrasound, as a function of lesion size, and that's not well known ever from mammography, especially for the smaller size lesions.

So I think the best we can do and this is why it keeps coming back to the best you can do at the current time, you can take the estimates

that you have and that's the purpose of sensitivity analysis, right, we can say what is the best estimate, what if it got slightly worse, what if it got worse than that, how would that change the management or outcome of patients?

I encourage all of you to read that Breast Cancer Research and Treatment paper by Allen, et al., because that's exactly what it does. It doesn't say here are the values and we know them. It says what happens when we vary these values, what is still the benefit or outcome for these patients? And that's all we can do at

00186

the current time, because to do these trials head to head to answer these questions will be another five, seven, eight years of data collection, especially in a cancer in a screening population where there's low incidence, and during that time you do, I think, a disservice to the women that currently have a need for the test.

DR. PAPATHEOFANIS: Thank you. Dr. Abrams, as the only oncologist sitting on this panel, can you share your thoughts on this subissue of dense breast?

DR. ABRAMS: I'm not sure an oncologist is the one to answer a screening question. I think the screening issue is complicated because it's not one of the pieces of information that we really reviewed. I think when I read the report and it was pointed out to me that they specifically didn't have data on these indeterminate cases where the -- so that's why they went with the larger palpable, larger tumors. And you know, when I first looked at that, I said well, if PET can't prove its role there, then it may not be able to prove its role in the others.

But thinking about that more, that may not necessarily follow. I think we still need the

00187

data in these indeterminate cases, which maybe they're indeterminate because mammography does

depend on density, and that may be an area where PET would have a true advantage as we've heard, because it gives it signal another way. But I don't, no data was presented on that so it's hard to have an opinion today other than what was talked about by the public comments.

So, I think the other thing is, we made mammography prove itself in screening by doing mammogram studies that took many many years to prove that they actually hopefully would save lives and bring some benefit, because there are some costs to biopsies, and anxiety, and all the issues that people who lived through the mammography debates know about. So I suspect, just speaking to the screening issue, other techniques that want to enter this arena as screening tools will have to go through that kind of testing also, and that at least wasn't presented so far.

DR. PAPATHEOFANIS: Dr. Guyton.

DR. GUYTON: I think another thing to is that there are biopsies and there are biopsies. There are needle biopsies, there are core biopsies

00188

and there are excisional biopsies. And to consider using the PET scan on a palpable mass is for a surgeon an anathema. When you can stick a needle into the thing, stick a core needle into the thing and find out what it is, you don't have to depend on its glucose metabolism. So that some of those issues come into evaluating these questions.

I think the other thing that can come out of the discussion today is to try to determine what data is needed by HCFA in order to make some of these determinations and that they can then determine what they need, how they might be able to go about it, as they have done with the national emphysema treatment trial, and arrange for those studies to be done. Study PET versus biopsy for nonpalpable mammographic abnormalities. Study PET versus mammography and ultrasound in DY

dense breasts or as identified in problem mammograms. Study PET versus present methods of determining locoregional disease after a positive biopsy, as Dr. Rollo suggested. Study PET on some schedule versus short interval mammography on follow-up with low or medium suspicious findings on mammography. Compare PET to sentinel node

00189

biopsy and axillary lymph node dissection in determining locoregional staging.

Those are ways of going about getting the information that is needed to answer these first three questions.

DR. PAPATHEOFANIS: I think that's well said. Dr. Flamm.

DR. FLAMM: Just to add one more piece to the discussion about other imaging choices and dense breast, I think there are a couple of other technologies that are being applied to looking at dense breast. You mentioned ultrasound, and MRI as well, both which function on the basis of different physical mechanisms for obtaining their signal than radiographic x-ray density, ultrasound characteristics, and MR is proton signal density. So I think that both of those technologies would need to be kind of at least put into the discussion in thinking about meeting this unmet need, where mammography is very limited in the dense breast patient.

DR. PAPATHEOFANIS: Anyone else on the panel that might want to add or subtract something?

MR. KLEIN: Just in terms of some data

00190

on this, there is a lot of data about what is missed and the percentage of misses in traditional mammography. And you know, I worked at Variant a number of years and we spent a lot of time taking a look at what cancers are missed. And in the breast cancer area it's very clear in all the data that anywhere between 70 to 82 percent, so that 80

percent of the cancers that are there are caught during traditional mammographic review, which is another way of saying that 20 percent are missed and are missed for a variety of reasons, either due to radiologic oversight, you know, very busy departments, they are on the images but they are just not picked up.

But in those cases, where 20 percent are missed, a third of those cases are in dense breast tissues, so we're looking at 7 percent, or seven out of a hundred times when it's missed, or seven out of a hundred mammograms will be missed, and they will be missed because of dense breast tissue. Now whether or not this is the best modality or not to detect that is a subject for discussion, because I think there were a lot of other points made also about ways in which cancers are missed either because there has been biopsy or

00191

because there's been breast augmentation or because there may be other risk factors, genetic risk factors, family history, whatever, that would be important.

But I think the dense breast issue is one area. But the reality is that seven out of a hundred will be missed, 20 out of a hundred will be missed and seven out of those will be because of dense breast issues, and some may even argue that that's a conservative number.

DR. GAMBHIR: I think that's right, but if you then go to Wolf grade DY -- that's actually across all Wolf grades, but if you now focus on the model where we are talking about the worst ones, or the highest density, it will be actually almost double that number, because those are the ones that mammography does even worse on, so yeah, I think there are real misses in these women that have to be addressed through PET and/or additional technologies.

DR. TUNIS: Sam, do you know much, or Dr. Flamm, about the performance of ultrasound or these other modalities that were mentioned in

terms of the these DY 4 breast densities?
DR. GAMBHIR: Again, the problem I

00192

think lies in that with the other technology as well, there's not good published data on a head to head comparison. There are studies underway now at several institutions that are looking at dense breast women with high risks, that is a family history in addition to dense breasts, where they are looking at MR imaging, ultrasound, mammography, and in some they are adding PET. Until those data come out, I can't give you a head to head comparison of the two.

I do want to say though, that from my other hat, which is more as a molecular cell biologist, what we're talking about sounds so primitive in that it's to me, just to put it in contrast, I raised this analogy the last time six months ago, that it's like saying prove to us that what applies in an x-ray on the left pinky applies on the right pinky, because you haven't proven it for the right pinky. To me it's not just a breast cancer issue, it's the fundamental biochemistry of these tumors. This is not the tissue it originated in. When we go later to the issue of recurrence, looking for staging, it's not which metastasis is present in the liver, where it came from, it's the fact that it's in the liver. We

00193

are limited in its size for sensitivity, and its specificity it determined by issues of inflammatory response and other background activity that's not anything to do with the origin of the tissue type. So when we look at these other categories, we have to be careful not to say oh well, show it to me in the breast literature. When we look back a decade from now later, you will hear in your own minds echoing these words, that that doesn't matter, it just doesn't matter that it originated from breast.

DR. TUNIS: Sam, what is it that

explains the 10 to 20 percent false negative rate for the axillary use in breast cancer, or, I don't remember exactly what the false negative was, but given that these tumors do consume 20 times more glucose or whatever, what accounts for a false negative?

DR. GAMBHIR: I think that's a very good question, and it applies to all cancers, not just breast. The main reason for false negatives tend to be, one, tumor burden at that site. None of these imaging technologies are looking at a single cell or a hundred cells or a thousand cells. You have to approach hundreds of thousands

00194

to millions of cells in a given site. We would love to have a technology that identifies these molecular areas when you're down to just one or two cells having that error. These technologies don't do that, so the smaller the tumor is, the smaller the lymph node metastasis is, the less our chance of being able to catch it on any technology, including PET. So that produces false negatives.

Then there are different degrees of FDG uptake by different tumor types. Not all breast cancers are absolutely equal. Ductal carcinoma in cyto will not be as metabolically active as infiltrating ductal cancer. Infiltrating ductal cancer tends to be a little more active than lobular, so different tumor types do have a range of glucose metabolism, and that also causes us to miss certain tumors, but both those lead to less than perfect sensitivity and again, then, it depends on not the origin of the tumor but the tumor burden at a given site. So whether it's lung cancer that has made its way into the axilla or whether it's breast cancer that's made its way into the axilla, it's the number of those cells at a given site that matters and the rate of glucose

00195

metabolism.

And the contrary is the specificity issue. It's not which tumor metastasized to the liver, it's what are the things that cause false positives in the liver or false positives near the bowel wall. It's not the site it came from. So although we can close our eyes and say no, no, but let's focus on the breast literature, really what we should be focusing on is for all these different tumors coming to this site, what's our probability of catching it at this site and what's our probability of being falsely positive. And that's the arguments that you know that I used at the last meeting across all those other cancers, and that's I think the more important way to look at this data.

DR. PAPATHEOFANIS: Thank you. Dr. Phelps and then Dr. Weinberg, and we are going to close this discussion.

DR. PHELPS: Just a brief comment. You know, still the issue with MR, you're switching from electron density to proton density, or hydrogen density, so it's still categorically a different issue. It's still the issue of gross density, and there is no relationship proven

00196

between disease and electron density or proton density. You know, and that's the point where we're trying to get everybody to come over to the other side to look at biology where there is fundamental proof in the relationship between biological process and disease, and then just take that evidence over to the patient with PET.

And it's not an issue of the value of x-ray techniques or CT or MR, we all know they are valuable, but it is to separate these two categories when we are trying to define the type of information that we are looking at and how we use that.

DR. PAPATHEOFANIS: Thank you. Dr. Weinberg.

DR. WEINBERG: Yes. I would like to perhaps assist Dr. Tunis in his question as to the

possible reasons for false negatives in PET. In nuclear medicine, size does matter, and just as Dr. Gambhir pointed out, you can miss large cancers and even in patients with palpable cancers, as Dr. Guyton has focused on, it's very helpful for some surgeons to be able to see whether there is multifocality associated with those large cancers, and that is again, a size

00197

question.

A technology was developed in Dr. Phelps' lab that's being used currently in animal imaging where you get one millimeter resolution. We have looked in protocol at patients who were injected with FDG, had the specimens removed, and we looked at those core specimens. You could see minute amounts of cancer. People have shown with autoradiography they can detect as few as 10 cancer cells, so PET is the heir to radiography, it really has a lot of power in terms of being able to see not only the large cancers but also very minute cancers, and so there's a lot of promise in this technology.

DR. PAPATHEOFANIS: Thank you. Donna?

MS. NOVAK: It seems like there's a spectrum of you know, from initial screening through, you know, we know we have a tumor and it's quite large. And it seems to me that these questions start in the middle of that spectrum somewhere and do not include the initial screening, so that really isn't part of our charge if our charge is in fact these five questions, and I think a lot of the discussion is really around initial screening.

00198

DR. PAPATHEOFANIS: So far, right.

DR. BURKEN: And that has to do with the fact that there are statutory reasons for that maneuver and the questions did start there because of a statutory exclusion of screening, except for mandated reasons such as mammography. But on the

flip side of that, I think this has been a provocative discussion on dense breasts and just because we don't have a question on the page and we may not vote on it today doesn't mean we will leave it behind.

MS. NOVAK: That's another interesting point. Can this panel say, you know, we voted on these five and this is our vote, and here is something else that we would have liked to have considered or want to consider in the future?

DR. PAPATHEOFANIS: Well, I think we have said that, and I just wanted to close the discussion on the dense breast by asking Sean if he has captured enough information at this point, since it's not an issue that we will vote on, and we're just going to move on from here. Is there any further discussion you would like us to consider?

DR. TUNIS: My only measure is as long

00199

as everybody on the panel feels like they've had their say on this issue for the record and for our consideration, that's the only measure of whether there has been enough discussion. So I don't know if anyone who hasn't weighed in on this wants to weigh in. As I said, we won't vote on it formally, but obviously all of this discussion becomes part of our internal consideration.

MS. NOVAK: I'll say that I think dense breast is a specific example, but I think initial screening in general as far as what gives better diagnostic help. One thing with mammography, if you haven't had a mammogram obviously, you have to wait a period of time. And so, I think there are other issues besides just this one, which I think is an example of an initial screening issue.

DR. PAPATHEOFANIS: Right. And that's not to say that we will not be charged with addressing that issue at a future panel meeting.

So with that, let's go would to what we do have, and that is the charge of working our way through these five questions and offering our

recommendations to Sean and to HCFA.
DR. PHELPS: Can I ask one thing?
DR. PAPATHEOFANIS: Go ahead.

00200

DR. PHELPS: So I guess the question I have asked, you decided against, about the -- I mean, I raised the issue that dense breast subpopulation is an issue of palpation and x-ray techniques, it is not a subpopulation in PET, so you're rejecting the including dense breasts in the general diagnostic population in question with PET?
DR. PAPATHEOFANIS: That's what we're going to do, that's the sense from this panel, and if it does come up again from the Agency, we'll look at it in that light. I think what you're seeing is we sort of have one hand tied behind our packs in that dense breasts means something to a lot of people and the data weren't cut that way, it's not to say the data don't exist, but it's sort of an 11th hour request when what we have been dealing with are these five questions. It's not an excuse, it's just that it's sort of a destabilization of what we can do at this point, and I think that's why I'm offering that perhaps we will look at this at a future panel meeting.
DR. TUNIS: Yeah, and I don't think the -- I mean, we have obviously heard from this panel that a number of these panelists consider

00201

this an extremely relevant and important issue, and so that becomes part of our deliberation in the 60 days or whatever from the time we get our Executive Committee ratification of whatever is decided out of this meeting, so it's not as though this closes off the conversation on the dense breast issues. So I don't know that that constitutes in your view rejecting your proposal or not, but that's not what's intended. We are not going to vote on your proposal.
DR. PHELPS: I think you rejected it

for the vote today.
DR. PAPATHEOFANIS: We did, yeah.
Well, I started this by reading question one, that was about an hour ago. Let's try it again. Question one. Is there adequate evidence that PET can improve health outcomes when used to decide whether to perform a biopsy in patients with an abnormal mammogram or palpable mass? I think what I would like to do is discuss this question, and I think maybe take a vote after we discuss this question so it's fresh on our minds and do the same for the remaining five.
So with that, any comments on question number one? Dr. Flamm?

00202

DR. FLAMM: One framework to begin breaking down this question is to look at what we know about the diagnostic performance of PET in this indication, think about how it seems to change management, and then think about the balance tables that were presented in terms of benefits and harms and thinking about whether PET improves health outcomes.
And one point I think is helpful in this indication and it also applies to the second indication, we had a fair number of studies estimating diagnostic performance of PET here, and while adding in abstracts may increase the end, it's reassuring to see that the diagnostic performance estimates coming out in the abstract literature are in line with what we know now, so I think, I personally feel like we have some sense of how PET performs in the patient population that was studied, and I'm referring specifically to the segment of the population that we have.
And then you go to the next step and think about the balance of benefits and harms, and I think it's of concern here that a patient using PET to avoid biopsy faces such a relatively high false negative rate of having a cancer not picked

00203

up by avoiding a biopsy. So I think the problem in question one is not so much the diagnostic performance data, and I don't think that bringing in the abstracts would change my mind at all about this, in this indication. It really is, and given that level of performance, how it would be used in this clinical circumstance, the net effect wouldn't help the population of patients.

DR. PAPATHEOFANIS: Okay. Anyone else?

DR. GUYTON: I guess I would agree with that assessment, and particularly the palpable masses, because they ought to be biopsied and are easily biopsied and then can be evaluated from there. And then if you then take a subsegment of the abnormal mammograms, I think there is a standard of care that's present at this time for treating that situation and it's not clear to me that adding PET to that standard of care is going to change the outcomes.

DR. PAPATHEOFANIS: Mike.

DR. MANYAK: I think that, you know, we wrestled with this issue of trying to be a little more inclusive with more data from the abstracts, and I mean, I agree with the strict criteria that has been used. We wrestle with this in our

00204

specialties along the same lines, and you kind of really have to go with something that's, in my opinion, critically looked at like this.

However, even, let's say we did accept that data, and I think there was some valid points about incorporating a lot of that data, it still doesn't answer that issue of the small lesion or the indeterminate mammogram and if it did, then I would say that would be an important point to consider here. But it doesn't change, so adding another thousand patients doesn't change the conclusions of question one, and I think that's an important thing to remember here.

DR. PAPATHEOFANIS: And I think in your recommendations for PET forward, one of your suggestions was doing just that.

DR. GUYTON: Yeah, and HCFA can decide what information it wants, to design a study to garner that information, and then determine how large a study they want.
DR. PAPATHEOFANIS: Barbara, do you have anything you wanted to add?
DR. MCNEIL: No, I think the data is incomplete.
DR. PAPATHEOFANIS: Anyone else? Well,

00205

I need a motion, I guess. Janet.
MS. ANDERSON: At this time, the chairperson, Dr. Frank Papatheofanis will call for a motion and will ask the voting members to vote. We are going to vote on the first question which I will read, which is: Is there adequate evidence that PET can improve health outcomes when used to decide whether to perform a biopsy in patients with an abnormal mammogram or palpable mass? And what you're going to do is, we will start with the for, and just simply raise your hand until I tell you that I have you marked. How's that?
DR. PAPATHEOFANIS: I'm sorry, I didn't get that.
MS. ANDERSON: We will start with the members of the panel who are voting in the positive, voting for the question number one.
DR. PAPATHEOFANIS: We need someone to make the motion first.
MS. ANDERSON: Oh, I thought you made the motion.
DR. PAPATHEOFANIS: I can't.
DR. FLAMM: I move that we vote.
DR. MANYAK: Second.
DR. PAPATHEOFANIS: So the motion is

00206

the question, is everyone agreed on that one?
DR. GUYTON: So a positive vote is that there is adequate evidence, and a negative vote is that there is not adequate evidence?
DR. PAPATHEOFANIS: That will keep us

from having to rephrase the questions, right.
MS. ANDERSON: Those who are voting for? Those who are voting against? No one has abstained.
DR. PAPATHEOFANIS: Unanimous in the negative.
Okay, let's move on to question two, and I'll read that one.
MR. KLEIN: Can I ask a procedural question?
DR. PAPATHEOFANIS: Absolutely.
MR. KLEIN: Are we, are our votes in each of these areas going to be binary in the sense of yea or nay for each one of these, or is there a possibility to answer these questions yes or no under certain circumstances or for certain indications?
DR. PAPATHEOFANIS: Typically, a question answered no, correct me if I'm wrong, Sean, meets with a question from me, which I

00207

didn't do, and I apologize for that, as to why you voted no, and in that way, that information is entered into the transcripts. Would you like us to do that, Sean?
MR. KLEIN: I guess what I was getting at --
DR. PAPATHEOFANIS: Because that gives you a chance to say, well, I voted no, but this is why.
MR. KLEIN: I was really thinking about as we move forward in some of the other questions, there may be some points that, because some of them are very sweeping questions, there may be some points as we move forward, even in the next one that we have to deal with, where it may not be as simple as saying yes or no. The answer might be, if the motion stated it this way, I would say yes, or I make a motion that this is an indication for recurring cancers, or situations where a prior biopsy would be indicated. Can one move as such, or is one limited to make a motion that's

precisely duplicative of the question listed here?
DR. PAPATHEOFANIS: No, we will entertain motions in language that you propose, and either vote that language or not.

00208

DR. TUNIS: You can either make a motion to amend any of these questions, and have a vote on that, or you don't have to change the question, you can simply make commentary on your vote, which becomes part of the record and is as important as your vote itself. And that's even true for the nonvoting members who don't have a vote, they can still make a comment in relation to a vote, you know, even without being formally counted as part of the vote.
DR. PAPATHEOFANIS: So then before we go to question two, is there a comment you would like to make on question one?
MR. KLEIN: My comment is I will have a comment on the other questions.
DR. PAPATHEOFANIS: Anyone else that might want to make a comment?
All right. Let's go on to question two.
DR. BURKEN: I would like to make a comment on question one. I was wondering how many votes on the panel, you know, might want to comment on the risk-benefit ratios that were highlighted by David Samson in his presentation, whether that played a part in the decision making,

00209

because I think David highlighted those and I would just be curious how others were responding to David's remarks.
DR. MCNEIL: I actually thought Carole said that very nicely in her summary.
DR. BURKEN: Okay.
DR. PAPATHEOFANIS: Okay. Question number two: Is there adequate evidence that PET can improve health outcomes by leading to earlier and more accurate diagnosis of breast cancer

compared to short interval mammographic, vis-a-vis three to six months, follow-up in patients with low suspicion findings on mammography and other routine imaging procedures?

This is where you comment, Michael.

MR. KLEIN: Yeah. I think that there has been a pretty healthy introduction of some data on the floor by our presenters, indicating that if there has been an occurrence, and in fact there has been prior treatment either because of biopsy occurred, maybe making it difficult for a follow-up review, or if someone is on hormone replacement therapy even though, for Medicare purposes, one might normally suspect that there would be dense breast tissue but found because of

00210

hormone replacement therapy. Recurrent cancer is obviously, or one can argue that genetic predisposition, there are a couple of genetic factors that fairly conclusively lead to a higher percentage rate.

But I would say in the case of an already diagnosed cancer, to get an, that current mammographic procedures fall very much short in terms of the ability to detect anomalies or reoccurrences, particularly if there has been some treatment or if there's been breast augmentation, an open excisional biopsy or whatever. So I offer that as a comment in terms of one particular way one might want to consider PET as an indicator in certain circumstances.

DR. GUYTON: But you're talking about a situation where cancer has been diagnosed. This question does not address that at all.

MR. KLEIN: Well, you're talking about short interval mammographic follow-up for patients with low suspicion findings. I look at that as an indicator, and while there's later questions that may deal with people that have been treated, this is the case where there is clearly an individual in the high risk, the reason for the short

00211

interval treatment would either be because of prior cancer or because some risk factor has been determined. What are the other reasons for short interval, three to six month mammographic reviews?

DR. PAPATHEOFANIS: Dr. Flamm?

DR. FLAMM: I think there is a clinical scenario where a woman who is coming in for a screening mammogram has something a little questionable on one view, they don't see a definite mass on the other view, and they are a little unsure, they would like the woman to come back in three to six months for a repeat mammogram and maybe it will make itself a little clearer over time. That's the type of clinical quandary that I think is also captured in this group.

DR. PAPATHEOFANIS: I agree, picking up disease and tracking a patient who you're not sure of, whether or not there is disease.

MR. KLEIN: That's the intent of the question?

DR. PAPATHEOFANIS: Yeah. Any other comments? Any interest in changing the language?

DR. GUYTON: I don't know that enough data has been presented, I mean, essentially no data has been presented on this issue.

00212

MS. ANDERSON: Call for a motion.

DR. PAPATHEOFANIS: It's called for a motion.

MS. ANDERSON: Would someone like to move that we vote?

DR. LERNER: Yes.

DR. GUYTON: Second.

DR. PAPATHEOFANIS: Any discussion on that motion?

DR. ABRAMS: I'd just like to add, this again, this is not uncommon, this happens a lot, there's millions of women getting mammograms, so you would think that this is an area that if you have another test that might add to the adjunctive procedures to replace mammography, this is the

place where you could do a many thousand, nay hundred of thousand patient studies to see if PET would really add, and I guess I'm repeating Dr. Guyton's comment that that is what needed. This would be a great improvement in the field if you didn't have to tell people, go home and wait six months, you might have cancer, you might not, we can't tell you right now, so I think this is really important to do such studies.

DR. PHELPS: And who would pay for

00213

that?

MS. ANDERSON: This is what we are voting on: Is there adequate evidence that PET can improve health outcomes by leading to earlier and more accurate diagnosis of breast cancer compared to short interval mammographic follow-up in patients with low suspicion findings on mammography and other routine imaging procedures?

Those panelists who are voting for? Those panelists voting against?

We have a unanimous against.

DR. PAPATHEOFANIS: Would anyone like to provide any comments regarding their votes or should we just move on?

DR. GUYTON: I think the comment is basically what Jeff said, this is a ripe area for HCFA to decide what information they want and go get it.

DR. TUNIS: And I would just say in response to Dr. Phelps's comment, which wasn't particularly audible, about who would pay for such research, that I think after we're done voting with these questions, if this panel wanted to have some conversation about how they think this sort of research ought to be at least prioritized if

00214

not funded, that certainly the panel could have a conversation about that. I don't know if Dr. Phelps meant it as a rhetorical question, but he's asked the question of me before so I'm

passing it along to you.
DR. MANYAK: Is that appropriate for this panel? I was led to believe that we were generally not to discuss financial issues and those kinds of things, at least that's what I recall.
DR. TUNIS: I think there is some recommendations here about --
DR. MANYAK: Because there are other issues along that line that are very serious in this particular issue with PET scanning, very serious, but that's not our charge or our purview today.
DR. TUNIS: The purview is not to consider the cost of the technology in making the coverage recommendations. The issue has come up, it has been raised by several panel members about you know, the need, the priority of this sort of research. So I think, you know, at some level, wrestling some with that as a policy issue, given that it's raised in the context of this as a

00215

coverage issue can be discussed. I think that's different from --
DR. MANYAK: That's a different question than what he mentioned. Who's going to pay for it sounds to me like a cost consideration, as opposed to saying it should be a priority, that's a different issue.
DR. TUNIS: Exactly.
DR. PAPATHEOFANIS: It's also in keeping with I think one of the future roles of the Executive Committee as the identity, or the responsibility of the Executive Committee shifts to an even purer advisory capacity, one of the issues that the Executive Committee will deal with is prioritization of research needs. And to have our panel for example, pass that along to th EC would give further guidance to that committee and move things along. Jeff.
DR. LERNER: For the purposes of today, I guess I'm sort of a strict constructionist, and

having read the document from the Executive Committee, we're just -- you know, I think we are voting properly according to that document, but I am glad that you're opening side comments on overall policy issues because there are lots of

00216

them that come out. But for the moment, that document does say that we're not supposed to -- studies that haven't been done -- I'm trying to phrase it according to the actual language of that document, but there may be studies that haven't been done that may be difficult to do, or may be costly to do, but that doesn't mean that you know, we can't say, well, they ought to be done. But we have to vote on the current evidence and that's how I understand that document, so as a strict constructionist, yeah, I would like so see those studies done, but I think it's irrelevant at this point.

DR. PAPATHEOFANIS: Right, it's irrelevant but it's important information. What will happen as soon as we close today's panel meeting is that I along with Dr. McNeil will put together a summary of this meeting in very much a decision analytic format, and try to convey to the Executive Committee, and we're both on the Executive Committee, why this panel behaved the way that it did, and why it took the votes that it dia. But along with that narrative, we can add specific recommendations regarding policy and I think they would be met with favor by the EC in

00217

certain ways.

DR. KRUBSACK: Did the panel address -- this also says, if the evidence is inadequate or insufficient to draw conclusions, the panel will explain the reasons for its determination and also form a judgment about the possibility of developing better evidence and the potential benefits of obtaining better information, and it goes on to say what are common obstacles to not

having adequate information, and that includes technology is relatively new, costs of performing study is high, funding has not been available. I think all of these apply to the present situation, so I think this panel is charged by its own bible to form its own guidelines to address those issues.

DR. PAPATHEOFANIS: Yes, and I think as we get into questions three, four and five, that discussion becomes even more relevant, and we will probably draft language that takes that into account. Okay. Dr. Conti?

DR. CONTI: Could I ask a question to the question?

DR. PAPATHEOFANIS: Sure.

DR. CONTI: You asked also about

00218

restructuring the question itself, rephrasing the wording, and I might propose you to consider this for perhaps a future meeting, to take question number two and look at it in terms of something like this. Is there adequate evidence, et cetera, compared, to use PET leading to an earlier and mor accurate diagnosis of locally recurrent breast cancer compared to short interval mammographic follow-up in patients with equivocal findings on mammography? That perhaps could be a specifically addressed question from the literature and something that would be more directed towards the appropriate patient population I think we're going to be talking about.

DR. PAPATHEOFANIS: Thank you. All right, question three. Is there adequate evidence that PET improves health outcomes when used to decide whether to perform axillary lymph node dissection? If so, is a more detailed analysis of sentinel node biopsy versus PET as alternatives to axillary lymph node dissection necessary?

It's kind of a two-part question.

DR. GUYTON: Not necessarily.

DR. PAPATHEOFANIS: Any discussion?

DR. MCNEIL: Well, I think, Frank, that

00219

the analysis for three is very similar to the analysis that Carole made for question number one, so I would say ditto to what she said there.

DR. PAPATHEOFANIS: Okay. Any additional comment before I ask for that language? Okay.

MS. ANDERSON: Then we need a motion to vote on question number three.

DR. LERNER: So move.

DR. PAPATHEOFANIS: Is there any discussion before we vote?

DR. MANYAK: Second the motion.

DR. PAPATHEOFANIS: We have a second, and no discussion, so you can take the vote.

MS. ANDERSON: Okay. Those voting in favor of question three as it stands worded? Those voting against question three? Okay. We have six votes, it's unanimous against question three.

DR. PAPATHEOFANIS: Any comments about your voting? Anything else you want to add to the record? If not, let's go on to question four. Is there adequate evidence that PET improves health outcomes as either an adjunct to or replacement for --

00220

DR. BURKEN: Excuse me, I believe we need to go to the second part of question three? I'm sorry; that was only if yes to the first part.

DR. PAPATHEOFANIS: Is there adequate evidence that PET improves health outcomes as either an adjunct to or replacement for standard staging tests in detecting locoregional occurrence or distant metastases or recurrence?

Dr. Flamm?

DR. FLAMM: One comment that I think I want to make to help when we look at some of the studies that are presented in this evidence, patients were selected into the study by virtue of having had equivocal findings or problem scenarios

based on conventional staging tests including CT, MR in many cases, and PET was used in those settings and those studies do report sensitivities and specificities of PET and CT. But one caution I think is important to note is that that's not a prospective head to head comparison of CT versus PET in all unselected patients.

In this type of study population, we've taken out the easy diagnoses for CT and so, it's not logical to directly say that because the sensitivity of PET may be higher than PET in this

00221

type of a selected study setting that one is interchangeable for the other and you can expect this diagnostic performance to be the case in all patients.

DR. PAPATHEOFANIS: Yes, Dr. Abrams.

DR. ABRAMS: This question is the one that gives me personally the most difficulty, because I think we can all relate to some of these stories that we have heard about how PET has helped in certain difficult clinical circumstances, like brachial plexopathy versus soft tissue invasion, like bone metastases versus advanced degenerative disease, where certain of our other tests don't work all that well and we know that by long experience, and having another adjunctive test can be useful, although I can also see here how you know, as Dr. Wahl pointed out, we might have to wait a very very long time to have a series of a hundred patients that were properly, you know, that had a prospective study done.

So I think in some circumstances, you are forced to look at smaller pieces of evidence, 10, 15-patient studies that if they are fairly, if the evidence is fairly distinct and coming from experienced clinicians and radiologists, is pretty

00222

believable, and I myself am struggling with having, with getting much better evidence. And maybe others have some thoughts about that, but I

think those, that's what makes this a difficult question. It sounds like there is some evidence that it has helped people in difficult clinical circumstances.

DR. GUYTON: In the evidence that we can consider, there is expert testimony, and consideration of -- there is another term that they used here -- other relevant information including guidelines from professional societies and other expert bodies, et cetera, so that also is evidence.

DR. PAPATHEOFANIS: That's right.

DR. GUYTON: And we are the jury.

DR. PAPATHEOFANIS: Right.

MR. KLEIN: I just have a question and a comment. The question is, in other areas where PET is indicated, as it is for lung cancer detection, where one could argue similar systemic concerns about the spread of disease beyond the local area, where they might be nodal involvement, Sean, do you know what the coverage is on that?

The reason I raise that is because I

00223

think you can argue that if it's indicated for cancers in other areas for this specific reason, for use of detection of lesions systemically or recurrences in other areas, then you could make the argument that it could apply here as well, and I'm just wondering what the coverage is, if there is coverage for this particular indication in other areas.

DR. TUNIS: For the cancers that are currently covered as of the December decision memo, we decided there that if there was a clearly proven single indication within a cancer, that other uses within the same cancer would be covered, subject to a set of restrictions. One would be that there wasn't evidence that showed in fact that they were not useful for a particular clinical use, and the other provision was that conventional imaging can't have already answered the question that you would presumably be asking

with the PET scan.

So in other words, for the lung cancer example, this would be covered for lung cancer, as long as there was documentation by the ordering physician that the treatment decision would be changed based on the result potentially, based on

00224

the results of the scan. Does that answer your question?

MR. KLEIN: Yes. Let me provide some useful background. I guess the problem I have may be similar to yours, Jeff, in that we could wait a long time to get data on this one, but it seems both intuitively and beyond intuitively proven with some concrete degree of comfort that if there has been a recurrence, that the regional or systemic involvement is not adequately answered by using imaging technologies, particularly as we have begun to start grasping how we're looking at this, which is in a more biological way. And the anatomical sort of spatial relationship between the tissue model that we have used is not really adequate in looking at the staging of disease, and I found some of the images pretty compelling, and I see other images as well that are even more compelling.

I've also seen the statistics, in fact this is a well established statistic, that when you find a cancer, if you go back the prior year, and two-thirds, 66 percent of the time, you will find that cancer one year earlier, and 50 percent of the time you will find it two years earlier.

00225

So it clearly means that our ability to detect cancers is not only lacking, but the ability to find the cancers in all the areas that they might be as they spread, as you get nodal involvement or further metastatic spread, is currently very limited.

So, in this one, whatever the vote is, I would hope that if the vote is to the no, which

would mean that there may not be adequate evidence, that I think we could at least establish for the record that there is some indication of such, of evidence, and perhaps there needs to be some further documentation to the point. But I'm not comfortable dismissing this point outright, because it's very clear that there is a propensity of evidence in the clinical setting, and while we wait to get the data, there are going to be a lot of people that will be misdiagnosed and will be lost.

And I think Kim Pierce made the point as a survivor and she is one of thousands of those who might benefit. So what I would be arguing for here is that there be some motion along the lines, if the argument is no, that there be some statement, there is indication requiring some

00226

further documentation to move to the category of adequate evidence.

DR. PAPATHEOFANIS: Dr. McNeil?

DR. MCNEIL: Like I guess Jeff, I'm the most conflicted about this particular indication. And as I'm thinking about it, I'm trying to think about it in terms of the data and the clinical consequences, and the feasibility of getting additional data as well as the problems with not getting additional data. I think we have to consider all four of those.

And as I listened to Rich Wahl, I was struck by one fact, which was that brachial plexus was an unusual situation, it occurred infrequently, 15 times in 8 years, but when it did occur, this was quite a dramatic way to diagnosis it, and there might not be other technology as good for that particular site of suspected recurrence. So that's, I think I could understand approving an indication that said suspected brachial recurrence, and maybe with a slightly broader mantel to that.

So then I get to the rest of the body, and I get in trouble and my logic, it's hard for

me to be clear about what's really going on here

00227

because I read the document, and the studies have the problem that Carole mentions in that the easy patients have been taken out of the pool, so that we're looking at only the tough ones, and even when we look at only the tough ones, there are some false positive rates here in several of the areas.

So I say okay, now what do we do? Then I say, maybe we take what Michael just said and we say we should do what was done for lung cancer and if all other efforts have failed, you go to this one. As I thought about that one, and I actually hadn't thought about it until you raised the issue, Michael, that one bothers me actually. And the reason it bothers me is that if we were to say downstream, this is really going to be a very powerful one-stop shopping for distant metastases in this disease, we have lost the opportunity to ever find that out by the approach that has just been suggested, because we will never get the data. We will always have the patients presorted by other modalities and then we will be left with the ones that were a problem.

DR. GUYTON: I don't see why a decision to allow that precludes us getting that data.

00228

DR. MCNEIL: Well, I'm just guessing that the radiology community is not going to rush to do that particular study. Now I could be wrong, but they --

DR. GUYTON: I don't see it.

DR. MCNEIL: Well, perhaps, but anyhow, if that were the case.

DR. GUYTON: I think you would find people would love to find, replace all the multiple scans with a single scan, given the same or better data.

DR. MCNEIL: Perhaps. I mean, I'm sure they would. The question is would they -- I don't

want to disagree with you, I'm just raising that as an issue. So what I come down on is that on the basis of just the anecdotal data that Rich Wahl presented, the locoregional, if that means brachial plexus, sounded pretty convincing to me. The other area does not sound convincing to me, and it looks to me as if it's begging for additional data. Now maybe I'm misinterpreting something in this document, but given the way the patients were selected, I'm not sure that I am, so I would just like a little help on my thinking here.

00229

DR. GUYTON: It sounds tome like the way the patients were selected was basically the way we're treating lung cancer coverage at this point in time. Is that approximately correct? If the findings are equivocal on the CT scans or whatever needs to be done, that's indeterminate findings in evaluation, that's how the patients were selected for the studies that have been presented.

DR. ABRAMS: If I understood what you said correctly, maybe it bears repeating one more time what they're doing in lung.

DR. TUNIS: The way the coverage policy is written is that if there could potentially be, HCFA -- right, if this was residual clinical uncertainty about appropriate management after conventional imaging, in other words, if the PET study may inform a change in the clinical management, that the PET scan would be covered. And the requirement is that the reason that it's being ordered is documented in the chart.

So whether that maps exactly to the scenario that you were describing that most of these studies are done in, is close. I'm not sure it's exactly the same, but it's close.

00230

DR. ABRAMS: But is that from metastatic evaluation in general in lung cancer,

or are we talking about pulmonary nodules and things like that?

DR. TUNIS: It's not specific to metastatic evaluation.

DR. ABRAMS: It's not specific to nodules, it could be any metastatic evaluation?

DR. TUNIS: Right, exactly.

DR. PAPATHEOFANIS: Dr. Wahl.

DR. WAHL: Since my name was mentioned, I thought I should just comment, and my intention wasn't to suggest that PET only had a role in imaging brachial plexopathy. Our experience is that is was uniquely superior to other methods in that particular setting and I couldn't convince our referring physicians to order any other tests.

But I would respectfully disagree with Dr. Flamm in how some of the studies were done in evaluating the comparative accuracy of PET. In bone scanning as an example, I think the study from Gary Cook as one, and having just reviewed this for the Seminars, was done as a prospective comparison, as I read it, between PET and bone scan for bone metastasis. And these were read

00231

independently, thus the bone scan wasn't used as a selector for the PET scan, and PET showed more lesions and had fewer false positives.

This was also true of the performance of PET in evaluating the skeleton in lung cancer, where it's now covered. So, you could easily argue as that point suggested, could PET replace the bone scan, and the answer would be yes. And I think several studies showed that where they were directly compared, and the entrance criteria were not an abnormal, or was not to be an abnormal conventional diagnostic imaging study, if I remember that correctly, and I think I do.

DR. MCNEIL: That's what it says here, Rich. Do you think that's wrong? That's what written in the table.

DR. WAHL: That is was -- well, the patients, as my understanding --

DR. MCNEIL: It says history of breast cancer, evidence of bone mets on bone scan in greater or equal to one other test.

DR. WAHL: My understanding, and I don't have the two papers with me, was that that paper and the study from Germany were done to directly compare patients with PET and bone scan,

00232

and some of the patients having normal studies.

DR. FLAMM: When you mentioned the Germany study, you've heard of the Bender study?

DR. WAHL: Yes.

DR. FLAMM: The Bender study specifically states that patients were selected on the basis of having equivocal or uncertain findings on the basis of conventional imaging, so I think that that is definitely a subselected group. I would link it specifically to the Cook study at that time.

DR. WAHL: I would have to review it to be absolutely certain, but Cook just wrote a chapter for a textbook I'm doing on PET and I did read -- well anyway, I believe that's how it was reported. I think the point is that PET appears to be able, even in difficult cases, appears to be able to find more abnormalities and be more certain about what they are than the conventional tests. I guess that would be the point.

DR. PAPATHEOFANIS: It's my intent to bring Dr. Larson back to the podium, but he seems to have volunteered.

DR. LARSON: I think that, I just wanted a point of clarification in the data that I

00233

presented. The problem of the 133 patients in that table that I gave you, Barbara, was with the clarification you'll see in the handwritten data, and again, I apologize for this, but you'll notice that actually based this categorization, which is six month follow-up as the gold standard with the available tests including biopsy and progression

on conventional testing, there really is quite a small rate of false positive. The problem is the false negative.

Actually, the bottom line where there was a positive PET with conservative management that was stable is the category of false positive, and that's only 7 out of 133 patients. The false negative group is significantly greater than that, and that's what accounts for the balance of the remainder of the inaccuracies. Remember, the accuracy here was 78 percent, so most of those were false negatives, so I just want to clarify that point for the thinking.

And again, this was a population that was selected because they were imaged because physicians referred these patients because the management was in question after conventional techniques were done. And this is actually a very

00234

important category and a very difficult patient group to manage, and I would submit that getting a significant fraction of an accurate management resolution, which my calculations suggest is about 78 percent, if that were followed, is very helpful.

But again, this is a very very selected subset, this is a group where the conventional techniques are equivocal.

DR. MCNEIL: So Steve, the false negative rate on this would be 7 -- I mean the false positive rate would be 7 over 7 plus 28, so it would be false positives over false positives plus true negatives right, so it would be about 20 percent? Do I have that right?

DR. LARSON: That's correct, but what I was thinking is the contribution to the inaccuracy in the whole population is quite small, but on the other hand, the false negative, the contribution, the thing that degrades the accuracy down to about 78 percent is primarily the false negatives.

DR. PAPATHEOFANIS: Thank you. I see that this Bender study obviously is one of the

pivots to this argument, and I wanted to ask Dr. Conti if he didn't mind coming up and giving

00235

the alternative interpretation to the data. There seems to be some issue with the data. Peter, are you still here? I would be curious to hear your interpretation of this study and why you think its placement and the way it was represented in the assessment might be less than right on.

And David Samson, is he here? You might want to power your laptop up again and let's take a look at the data for part three, which is I think where we're at, so we're all clear as to what we're talking about here and why we are forming these conclusions.

DR. TUNIS: While we're waiting for that, someone was nice enough to hand me the actual language from the coverage decision regarding how it's worded, so I can read it for folks if they are still --

DR. GUYTON: Please.

DR. TUNIS: So, for staging and/or restaging for the covered malignancies, PET is covered in clinical situations in which the stage of the cancer remains in doubt after completion of the standard diagnostic workup including conventional imaging, or the use of the PET could potentially replace one or more conventional

00236

imaging studies. And in addition to that criteria, the clinical management of the patient would have to differ depending on the stage of the cancer identified. In other words, the test would have to have made a difference. So the stage has to remain in doubt after conventional imaging or it's felt that the PET could replace conventional imaging, at least one or more studies, and the treatment would change as a result of the findings. So that's the way the current coverage decision is structured, so if you want to model this one on that one is open to discussion.

DR. GUYTON: The other issue about the question is it states, is there adequate evidence that PET improves health outcomes as either an adjunct to or replacement for standard staging tests in detecting locoregional recurrence or distant metastasis or recurrence. So if Barbara feels that it's a good test for locoregional disease in the shoulder, she has to say yes, and then we put conditions on it.

DR. TUNIS: She has to say yes or she has to amend the question.

DR. GUYTON: Right, but the way it's stated, she would need to say yes.

00237

DR. TUNIS: Right.

DR. PAPATHEOFANIS: Dr. Conti.

DR. CONTI: Just again to remind you of the ACR and SNM's position on this, representing nearly 50,000 practicing radiologists and nuclear medicine physicians, we would vote yes to this particular indication. Now, also, in terms of the issues on the Bender article, the things that I was concerned about, and perhaps I misheard them but I just wanted to make sure. Number one is that this article also evaluated patients under routine clinical conditions, so it's not the type of study perhaps that one might decide on performing in a prospective fashion, but it does reflect a clinical practice scenario which, being a country physician myself, I like to do that.

It talks about patients being followed up with -- excuse me -- who have been completely evaluated and followed up for at least six months, so clinical follow-up is a component of the verification process in this particular paper, which I think was not mentioned in the analysis. In particular, if you look on page 1689 of the article, only patients were included where results had been verified by histology, except for a few

00238

cases, four, where extensive disease was verified

by clinical course. So in fact, there was reasonable criteria established and used to establish whether or not there was a disease in the location of interest.

The other point I wanted to make was that these patients all were part of a routine workup for staging, usually consisting of a physical examination, axillary lymph node ultrasonography, thoracic abdominal CT and/or MRI, bone synthegraphy, and serum tumor markers, so all the patients had a regimen of routine tests in addition to the PET scan. So they weren't screened out on the basis of a particular finding on a routine test, they were all studied with the technologies.

The PET scans were later independently compared to the standard imaging, so they were rereviewed and compared independently to the original performance of the study. Those are my comments.

DR. TUNIS: Dr. Conti, I just want to ask one question. Does your society develop any sort of professional, do you have a formal process for doing clinical guideline development for the

00239

nuclear medicine community?

DR. CONTI: Yes, the Society of Nuclear Medicine does.

DR. TUNIS: And is this a topic that --

DR. CONTI: Yes. I was corrected, ACR does also.

DR. TUNIS: And has any been issued on this topic, use of PET for breast cancer?

DR. CONTIN: For use of PET, yes, not for use with breast cancer, in other words, use of PET across the board.

DR. TUNIS: Okay.

DR. MCNEIL: Could I ask one more question while you're there? I want to make sure I understand this article correctly. Do you have any idea why only 63 of the patients ended up having CT and 75 ended up having PET?

DR. CONTI: There was some MR done instead of CT.
DR. MCNEIL: No.
DR. CONTI: It says CT and/or MR.
DR. MCNEIL: Yeah, but if you add them up, it comes out to 63, unless this table, unless table 12 is wrong, there is a dropout of 12 patients between taking CT and/or MR and PET.

00240

DR. CONTI: They do say that this was an optional examination in their methods section, so I can't explain why the authors chose to do that.
DR. MCNEIL: Right. Just one more point to make sure I have this right. You disagree with this notation that the PET was not read blind, is that what you just said?
DR. CONTI: Well, I'm just reading what the article said.
DR. MCNEIL: That's why I'm asking you, you're the only one with it in his hands.
DR. CONTI: It says quote-unquote, PET results were later independently compared to standard imaging modalities, x-ray, CT/MR, ultrasound, mammography, film synthegraphy, quote-unquote.
MR. SAMSON: And I would like to clarify my point of view on this. Later on in that same paragraph on page 1689, it says, only patients were included, and this is I think a translational error, where results had been verified by histology except for a few cases. So that I read as meaning they had histologic confirmation as the reference standard for 71 out

00241

of 75 patients, and they used follow-up in four cases, and that's, that was the fundamental criticism I had with the Bender paper. It didn't seem to make sense to me that they could do histology in 71 patients for a number of different sites, they did bone, lymph nodes, local sites,

liver, it doesn't seem logical that they would be sampling lots of negative sites in all 75 patients, and there just isn't enough detail to really know what the reference standard was for all sites for all patients, and I think that's the kind of detail we should demand of studies like this.

And then also, it says PET results were independently compared to standard imaging modalities and names them, but that's not the same thing as saying PETs were read blindly to the reference standard because what is the reference standard, it's not really clear.

DR. CONTI: Again, all the patients have been verified, either histopathologically or by clinical follow-up, so we know that they have disease or not disease. So the issue is, we're using standard radiological procedures which we rely on every day in clinical practice to

00242

determine the presence or absence of this disease, and you're telling me perhaps that that's not a reliable source to compare the PET imaging data to, and I --

MR. SAMSON: That's not what I'm saying. What I'm saying is it's not clear from this article whether the reference standard for the sites that they were assessing, the recurrences of different anatomic locations, whether the reference standard was histologic of whether it was clinical follow-up, this paper is not clear on that.

DR. CONTI: Again, I'm sorry to be argumentative but the fact is, it says four cases were not histologically confirmed, they used clinical follow-up on the patients that were evaluated, so I'm not sure I understand what you're talking about here. And let's also keep in mind that with metastatic disease, we are not going to be able to biopsy every particular site, as we talked about earlier.

MR. SAMSON: No. And I think it's

perfectly legitimate to use follow-up as a reference standard, and I made that point in the presentation this morning. What I'm saying in

00243

this particular article, we don't know what the reference standard was. It's not clear.

DR. PAPATHEOFANIS: Any additional data on this part three of your assessment that you want to comment on?

MR. SAMSON: The only other thing I would mention is that if you want to take a separate look at the issue of locoregional recurrence and especially at the brachial plexus, we have one published study by Hathaway that looked at issue in 10 patients. I think lots of other comments have been made about how PET may be particularly useful for this particular indication, but I think this is a pretty small evidence base to make that kind of conclusion.

DR. PAPATHEOFANIS: Great, that's very helpful. Dr. Flamm?

DR. FLAMM: As long as we're pinpointing details of language in the paper, let me just clarify two things. One may help Barbara in your initial question about the number of CT patients that are good PET patients. It says all patients were part of a routine workup for staging usually consisting of physical exam, axillary lymph node, CT, da, da, da, so they may have had

00244

some but not necessarily all of those things.

And the second point is the second to last sentence in that same paragraph after what Dr. Conti read, it says patients were referred in order to confirm or dismiss a suspicion of tumor recurrence or systemic disease, or distant metastasis in undecided/equivocal cases, so that's where I was getting that from.

MR. KLEIN: Just a question. Sean, I know you had a summary of the data. You wouldn't happen to know what the burden of proof was in

presenting, in getting that indication in terms of what led up to those conclusions, or those indications?

DR. TUNIS: For the lung descriptors, you mean which was the indication that was considered the prove indication?

MR. KLEIN: Yeah, the one you read that was indicated for, I'm just wondering what the burden of clinical efficacy date, the data for efficacy was to produce that result.

DR. TUNIS: Mitch, do you want to talk about that at all, in terms of the December 15th memo? I guess on lung cancer is what you are exploring. See, in lung cancer we had a covered

00245

indication for the pulmonary nodules, if I recall; is that right?

SPEAKER: And initial staging.

DR. BURKEN: That's correct, for the evaluation of solitary pulmonary nodules and also for staging nonsmall cell lung carcinoma. But as I said, you know, many of you are familiar with, I think just about everybody in the room is familiar with the December 15th decision memorandum where we extended coverage to many other indications by tumor type as long as there wasn't a particular contraindication.

DR. TUNIS: But I guess they're asking what sort of studies did we have for the staging, restaging in lung cancer, for nonsmall cell, how do those compare to these sort of studies we're looking at here, like are these studies worse, or better or about the same?

DR. BURKEN: Unfortunately my memory fails me, but there was a fairly good British study that really helped us to get into the particular area for lung cancer. And let me see, I'm not sure it's going to be in my folder her, in fact I'm positive it's not going to be in my folder, but there was a particular article that we

00246

used as evidence for lung, for extending that.
DR. MCNEIL: Was it better than these data?
DR. BURKEN: I remember the study, I don't have that particular study in front of me. I didn't think we would be getting into that particular issue and I didn't bring all my PET material with me, but I have several notebooks worth of PET articles back at the office, but I remember being, you know, I'm not being very scientific here, but I remember it being a fairly good study, certainly strong enough to go to bat with.
DR. PAPATHEOFANIS: Dr. Conti.
DR. CONTI: Just to fill in the gap perhaps with the CT issue that we talked about in the Bender study. If you look at the Huebner study, he also looked at CT versus PET and in my document from ACR and SNM, I did quote those numbers and again for the record, the sensitivity in the 57 patients that had PET scan, the sensitivity was 85 and specificity 73 percent, compared to CT that was only done in 44 of those patients, the numbers were 71 and 54 percent. So at least you have additional data to show that PET

00247

is superior to CT with regard to detecting metastatic disease.
DR. PAPATHEOFANIS: Thank you. Any other comments. I think we have almost gone all the way back to -- sure, Dr. Phelps.
DR. PHELPS: You know, I think when we look at the TEC assessment criteria, even on Sam's papers we read the methods of the Blue TEC assessment, and we understand why that criteria was used. But on the other hand, when you go just to those strict and rigid criteria, you're setting a weight to all other information of zero, whether we recognize it or not. When we start looking at other evidence, you start shifting back to that. So all of the thing, you say the value that it provides is zero, and we know that that's not

true. And so, you know, maybe it shouldn't be equal to the peer studies, but it does have value, so it should have some weighted value in the decision that you make.

You know, also in the real world, where patients are being taken care of and you're doing research, you know, things are not so easy to build large populations in these criteria, so that's the real world we live in, and its weight

00248

shouldn't be zero.

DR. TUNIS: The only way I disagree with that, Mike, is for purposes of a TEC assessment what we're trying to do there is formally summarize the kind of better half of the rigorous scientific literature so at least we know what the rigorous side of the world has to say about this stuff. The reason we don't set the other stuff to zero is that we have meetings like this where, you know, Dr. Wahl can talk about his experience in Vancouver and Dr. Conti can talk about additional studies, and so that information is making its way into the considerations of this committee through all kinds of avenues other than being summarized in the TEC assessment.

So I think that for purposes of the TEC assessment, we're trying to summarize the more reliable body of scientific literature, and the rest of this meeting is about bringing all that other information forward, maybe not in as systematic a fashion, but it's not systematic information. So I don't think it's true that it's set to zero, I think it's just coming through in a different form.

DR. PAPATHEOFANIS: Dr. Conti.

00249

DR. CONTI: One last very quick comment. Just keep in mind that we recommended again, that this be at physician discretion. We would implore you to consider physician discretion in determining whether or not a patient needs

additional studies to make a diagnosis. Your question also poses as an adjunct, which I also think you should seriously consider the use of that particular word in your decision.

DR. PAPATHEOFANIS: Thank you. I was going to say, we have almost gone all the way back to Jeff Lerner's question about quality assurance and so forth, and this little side bar illustrates the process that we went through in choosing one of the data points if you will, of our consideration, and I'm pretty convinced that in the document that David and Carole prepared, due diligence was done, and I think it's a fair representation of the information that was there and I haven't heard otherwise, so again, for what that's worth, I commend them.

We've got number four on the table still, and we've got the data in the TEC assessment, we've got the data from public commentary, we've got individuals here with their

00250

own personal experience. Anyone else want to add to this discussion?

DR. GAMBHIR: (Inaudible) data, if you go and add the abstracts in this category, that doubles the N, okay? So it's like saying instead of Bender is just 75 patients, there are an additional 75 there, the Huebner article doubles, and then there was now management percentages that have been noted, and those management percentages in the abstract show that after conventional imaging, by adding the PET, 30 to 40 percent of patients change management due to the PET. So I think that data has to be weighed, and that changes managements occurring because you have now been able to understand whether it's locoregional recurrence, axillary recurrence, and/or distant recurrence, and all those data then, even though they're not yet in publication form, need to be weighed into the vote you're about to make.

DR. BURKEN: I would, you know, note some caution with respect to abstracts, although

you know, certainly much valuable information is in abstracts, there hasn't been a chance to really review the methodology and look and go through it carefully to see whether there are certain types

00251

of biases in those studies. So I think certainly there can be a lot of good information available in abstracts but I think there has to be some caution as well.

DR. PAPATHEOFANIS: Barbara.

DR. MCNEIL: Frank, I am confused beyond belief about what to do in this one, and I guess I'll just throw out some thoughts, and they may not be right, but I'll just throw them out. The first one is, in rereading this table, Hathaway stands out and Rich Wahl's comments about brachial plexus stands out, so I kind of have a feel for that.

I also think that from a policy perspective consistency is good, so that if in fact these data on looking for metastatic disease were actually equivalent in quality to the same data that led to the decisions in December, that would influence my thinking a lot because I think when you're making policy, you want to have some sort of consistent framework for making those recommendations.

If these data are not the same or of lower quality than the data that went into the December 15 judgment, and if we look at these data

00252

as they stand, then I'm really troubled, I don't think they hold up. I just looked at the Huebner article and it's a retrospective study with all kinds of people dropping out.

So I don't really know what to do. I guess what I'm doing is asking for some kind of potential advice about how to split up this question in a friendly way before we go down a vote that may not be helpful.

DR. MANYAK: You know, maybe this is a

role for changing the wording of the question, because I have exactly the same conflict. I'm really torn with this, because there is clearly anecdotal evidence that suggests that PET is valued in a subset of patients, but it clearly does not meet the criteria of strict review. So I mean, which way do you want? And frankly, I'm uncomfortable just saying no to this outright, yet I think it's very important to adhere to the criteria that have been set up which are good criteria.

So I think maybe either we vote on this issue and then add a significant comment after, or we change the wording of the question. I think we have to do one or the other.

00253

DR. PAPATHEOFANIS: I would actually favor more of the latter, that we actually change the wording. Dr. Abrams?

DR. ABRAMS: You know, I thought what Dr. Gambhir said earlier about should you ignore the evidence from other diseases, I did feel like we should ignore it in the screening question, because I think you are dealing with different issues when talking about primary tumor. But I would take his point here that we shouldn't totally ignore what has been found in other metastatic diseases in terms of, you know, its ability to help with differential diagnosis.

And so, I don't view the lung data or the other indications as necessarily, you know, this data has to be as good as that data. In a way I view it that data sort of helps me here, because you know, we are dealing with somewhat similar issues, and I think as best we understand the biology of these metastases, there are some similarities. They may not be identical, but at least the principle that this test is operating under, it seems to make sense that that's the understanding.

So for me, that was why I was thinking

00254

that the wording that we just heard on that lung policy, posing it as an adjunctive as opposed to a replacement sort of makes pretty good sense. Replacement, I would have want to have better evidence; adjunctive, I think that's sort of where they went with the lung data, and this data speaks to that point too.

DR. PAPATHEOFANIS: Good. I want to call just one more person to the podium to get a little more insight, and that's Ed Coleman if he's still here. Dr. Coleman, share your thoughts on that proposed language change.

DR. COLEMAN: I'm Ed Coleman from Duke University, am a professor of radiology. I have received honoraria from GE, from Radiology Corporation of America, from other mobile PET vendors to give lectures. I have been doing PET scanning for many years now, starting back when I was a resident at (inaudible) Institute of Radiology. I have had one of the most active clinical PET centers at Duke. Over the last couple of years we have started doing more and more patients with breast cancer, and it's primarily in this indication that we're talking about here. And it's generally as an adjunct to

00255

the other imaging studies after they have been completed and they have indecisive conclusions based on the other imaging modalities.

So I think that putting it as an adjunct would be appropriate. I think that as e get more data, we're going to find that it does replace the other imaging modalities, and a wording similar to what's been used for staging of the malignancies in the December 15th memorandum would be appropriate for this use in breast cancer.

DR. PAPATHEOFANIS: Give us a sense of, and I know this is putting you in an awkward position, but let's say the breast cancer specialists at Duke, let's say the language comes

in, and it is an adjunctive test, is this going to open the floodgates, is there discretion, is there an understanding by breast cancer specialists of the appropriate use of PET? I mean, I'm trying to get a sense of where the real world stands.

DR. COLEMAN: I think that the oncologists are learning extremely rapidly how PET is best used in the management of their patients. They've learned a lot with the indications that we have now; with the expanded indication that's

00256

coming in July, certainly they will learn more, but I think that the oncologists are getting very savvy on how to best utilize PET in answering these specific questions to their patients. It's not just going to open the door that everybody that has breast cancer needs a PET scan. I think that it would be specific patients with specific questions as to does the patient have recurrent disease, metastatic disease, and will be used specifically with the other imaging modalities to answer that question.

DR. PAPATHEOFANIS: Great, thank you. Any questions for Dr. Coleman? Thank you.

Well, anyone good at word smithing or are we going to just change a couple words around? Barbara is very good with commas.

DR. MCNEIL: That's an inside joke.

DR. PAPATHEOFANIS: That's an inside joke.

DR. MCNEIL: All right, I'll try a word smithing, given what we've just said.

Is there adequate evidence that PET improves health outcomes as an adjunct to standard staging tests in detecting locoregional recurrence or distant metastases/recurrence when results from

00257

these other tests are inconclusive?

I think that's the spirit of what the lung cancer, and consistent with --

DR. MANYAK: Would it be inappropriate

to say anecdotal evidence, is that the --
DR. PAPATHEOFANIS: No, I think she meant results from other imaging tests.
DR. MANYAK: Right, but what I'm saying is we change the, instead of adequacy, anecdotal evidence?
(Chorus of nos.)
DR. MANYAK: That's what it is, folks, I really think, but you know, that's okay, we don't have to call it that.
DR. TUNIS: There was a consensus that that was a bad idea though.
(Laughter.)
DR. MANYAK: That's the first thing everybody agreed on today.
MR. KLEIN: Is it worth reading -- I think that was pretty good what Barbara put together -- is it worth reading the lung statement again, just in case there's a little trailer there that might be interesting to add?
DR. PAPATHEOFANIS: Well, you know, the

00258

lung is for the lung, and I think we want to move beyond that, because I think the language there was a little different than what we're hearing today, and I think we're pushing this as much as we can.
DR. TUNIS: I mean, just to respond, on the lung issue, we were careful to make sure that the approved indications in December, you know, met reasonable but at least minimum standards of scientific adequacy of evidence, so it wasn't a gimme or something like that. So you know, without being able to cite you how big the study was or what flaws it was, there was at least one good study in this area, and that clearly exceeded the margin of anecdotal evidence. Beyond that, I can't say much about the lung question, but I think this has to stand or fall on its own merit.
DR. PAPATHEOFANIS: I agree, so I think we should back off from the lung analogy because I think we've taken this as far as we can. And I

think that rather than use the word anecdotal, I think what we're trying to say and while we're trying to be consistent with previous policy, is that there is a certain level of data, there is some discrepancy in the interpretation of those

00259

published reports, there is a significant body of anecdotal information, and we're taking all of that into account in changing the language and voting on that.

So with that, if you want to reread -- I'm sorry, go ahead, Donna.

MS. NOVAK: I had a question. Because of the wording of inconclusive, does some of this evidence actually indicate that it's better? I guess maybe if it does not indicate more lesions, then you can say that's inconclusive and go to the next step. I'm just having a little problem because it seems like there's some evidence that it might be a better test.

DR. MCNEIL: I was using the word inconclusive with regard to the results of the tests.

MS. NOVAK: I understand that, right. I understand that, and I guess I wanted to make sure that there is enough leeway that if a physician felt it was a better test, that they could go on, even though there might be some --

DR. PAPATHEOFANIS: You mean skip the test in between that might turn out to be inconclusive?

00260

MS. NOVAK: It's an adjunct, so you can't skip the tests, but I guess I'm saying that -- maybe I'm convincing myself that if the original test doesn't show any additional lesions, we could say that's inconclusive because it didn't show anything, and go on to the next step.

DR. GUYTON: But if there's significant clinical suspicion, that would be the plan.

DR. NOVAK: I just needed to convince

myself that there would be some way a physician could order those tests if the first test they didn't accept, for whatever reason.

DR. PAPATHEOFANIS: Go ahead, Dr. Conti.

DR. CONTI: I think it's important clinically to understand that we may need to know tumor burden to make certain decisions in these patients, so you might have an equivocal finding that's on a bone scan or CT scan, and even if it is perhaps a solitary lesion, you might act differently than if you knew you had widespread metastatic disease. So I think you need to have enough flexibility in this indication to allow physician discretion, because that decision, the physician has in his mind a certain pathway that

00261

he or she is going to go down if they know certain pieces of information.

So you might have a test that has one lesion, but if they have two, they're disqualified perhaps from a particular protocol, so I think you -- make sure that we have enough flexibility so that the physicians ordering the tests have enough discretion to determine which tests, or which pathways to choose from.

DR. GUYTON: But you're talking to Sean at HCFA, you're not talking to us, because we have to make a decision based on the evidence that's in front of us.

DR. CONTI: The issue that I'm --

DR. GUYTON: But that's not the issue that we're talking about.

DR. CONTI: I'm talking about the word inconclusive. I just want to make sure we understand what the use of that word is, because inconclusive might mean that there is no evidence, it might mean that there is an equivocal finding, or might be the patient has widespread disease from some other process.

DR. MANYAK: But that's the definition of inconclusive, I believe, isn't it?

00262

DR CONTI: Indeterminate might be a better word rather than inconclusive because you might make a conclusion, or you might be inconclusive because you're not --

DR. MCNEIL: Well, maybe, I guess two comments. The Bender article that meets the criteria said undecided or equivocal, so maybe that would be more appropriate. But my guess is, I would vote against the motion as I just word smithed if it this were to be used to measure tumor burden. These data that are presented to us have nothing to do with tumor burden in a quantitative sense, they just had to do with sites of disease, so I think if you want to introduce that, then that should be put on the table as another question. If it gets rolled up into this one, you will change my way of thinking.

DR. CONTI: Well, we can forget the discussion of tumor burden, that's not a problem.

(Laughter.)

DR. MANYAK: Boy, did you scare him off.

DR. MCNEIL: Well, no. I think we have to read the data the way we've got it.

DR. MANYAK: I agree with you, I think

00263

you're absolutely right.

DR. TUNIS: On that point I guess I would ask Dr. Abrams. I mean, what I understand this motion to be about is that if the clinical information would potentially change the management strategy, treatment strategy, then that's relevant information. So if it's a solitary lesion versus ten lesions then we need to know that. You know, you as a clinician would know in breast cancer. I figure one lesion in the bone is enough, you don't need to know that there's ten, but I'm not an oncologist.

DR. ABRAMS: I would agree with you that if the CT scan gave me five metastatic

lesions in the liver and a PET scan gave me seven, I'm not sure that would help me very much, so I'm not sure I would need the PET scan in that circumstance. On the other hand, if I had elevated liver enzymes, couldn't find any other explanation, the CT was negative, maybe a PET scan would be indicated in that circumstance. So I mean, that's how clinicians will have to use this. And I agree with you, it should be to inform decision making.

And with that in mind, I just, you

00264

know, my interpretation of health outcomes included that. Some people use the word health outcomes as you know, end points of survival or disease free, those sorts of things. I included in health outcomes that it changes one's decision making and that may affect treatment choices which have their own morbidity, et cetera. So, I just wanted to make sure we were okay with health outcomes as well.

DR. PAPATHEOFANIS: Sure. Any more word smithing?

MS. ANDERSON: I'm going to go ahead and read what we have so we know what we're looking at and see if there's a word or two that we want to change.

Is there adequate evidence that PET improves health outcomes as adjunct to standard staging tests in detecting locoregional recurrence or distant metastases recurrence when results from other tests, and some did mention imaging may be placed in this area, are inconclusive? So it's either tests or imaging tests.

DR. MCNEIL: Just tests.

DR. PAPATHEOFANIS: Did you have a comment, Carole?

00265

DR. FLAMM: Well, I guess I'm just wondering if there's going to be a companion piece of what's left over after we've modified this.

Are we splitting this out basically into two different questions and votes? Is there going to be any specific discussion as a replacement for? That's kind of being silent then, if we change the language just to be a vote on PET as an adjunct to, there is something left over.

DR. GUYTON: I think the committee could decide to do that if they wanted to.

DR. PAPATHEOFANIS: What would help you, Sean?

DR. TUNIS: Well, it sounds like you dropped the replacement because the feeling was there may not be any evidence on that, so it probably would be useful to frame that as a question and then vote on it, since it is part of this question.

DR. PAPATHEOFANIS: So it's two questions then. One is replacement, and the other is the new one, the word adjunct. Okay. Any other word smithing? Let's start with the one where the words replacement are left in place.

MS. ANDERSON: Could I have a motion?

00266

DR. MANYAK: So move.

DR. MCNEIL: Could you read that again?

DR. GUYTON: Wait a minute. Why don't we stick with the one that we smithed?

DR. PAPATHEOFANIS: Well, we've smithed both really and created new ones, but okay, let's do that. We're going to go with the one that you created, Barbara.

DR. MCNEIL: So that the outcomes as an adjunct to, that one.

MS. ANDERSON: I need a motion to vote.

DR. MANYAK: So move.

DR. GUYTON: I will second, if you will read it again.

MS. ANDERSON: This is the question. Is there adequate evidence that PET improves health outcomes as adjunct to standard staging tests in detecting locoregional recurrence or distant metastases recurrence when results from

other tests are inconclusive? That's what we have.
DR. MANYAK: I think the wording was, when results from these tests are inconclusive.
DR. MCNEIL: From other.
(Inaudible colloquy.)

00267

DR. MANYAK: Leave other, okay.
DR. MCNEIL: Is there a value in having this second recurrence in here, distant metastases/recurrence, is that any value?
DR. GUYTON: Yes. It could be both, it could be several times recurrent disease.
DR. PAPATHEOFANIS: Any further comments on the language as it stands now?
MS. ANDERSON: Okay. We have the motion, so I'm just going to carry the motion and we will vote on the language that I just read.
DR. PAPATHEOFANIS: Well, on the motion itself.
MS. ANDERSON: Those voting for? We have five votes for. Those who are voting against? Those who are abstaining. We have five votes for and one abstention.
DR. PAPATHEOFANIS: Boy, would we want to know what why you abstained.
DR. LERNER: I am trying to cope with the burden of evidence. I guess I'm not comfortable and I guess I need to see more and in a sense, the people who voted yes said they wanted to --
DR. PAPATHEOFANIS: So your conflict is

00268

with the literature, the evidence in the literature.
DR. TUNIS: And it would on this one, it would help to know that Dr. Manyak had proposed the word anecdotal I think to reflect some sense within the conversation that while the evidence, while you just voted that the evidence was adequate, that my sense from this discussion was

that the committee felt that it was barely adequate or just adequate, and maybe that's what you're saying, Dr. Lerner. I just want to make sure that if anybody on the committee disagrees with that characterization, they can let us know now, just because that may, we would take that into account as we discuss this internally.

So even though there is no such thing as saying barely adequate, you have voted that it's adequate, but the sense I'm taking away from the conversation is that it sort of just got over the line, and if somebody disagrees with that on the committee, I would be interested in hearing that.

DR. GUYTON: Well yeah, I think it probably does meet a higher standard than that. Jeff was talking about it earlier when he said

00269

that the propensity of the clinical evidence that was presented was very positive for this particular indication, and I think we've heard very strongly from the people who are involved in the clinical activities related to this process that they are convinced themselves and they have convinced us that the evidence is adequate, and I think it's more than just barely adequate.

DR. MCNEIL: Sean, I actually think it's barely adequate and I think that I voted yes for this, but I voted because it just hit the line, but if we were to have other studies like this, with this level of evidence, I'm not sure I would vote yes again. I mean, I think this was kind of a, it was that close to me and on another day, if I woke up on the wrong side of the bed, I just might not be able to vote yes with this level of evidence.

DR. TUNIS: Good thing you flew first class.

(Laughter.)

DR. MANYAK: I would also like to add that I believe it was barely adequate. I mean, the dust cleared and the runner was safe at the

plate, and that's really the way I looked at it.

00270

It was slightly over the line, enough to convince me after this discussion, but I still like anecdotal even though you guys don't like it, but in the spirit of moving forward, I will desist from any further discussion.

DR. FLAMM: I would also agree that it was a very borderline decision for me, and that one of the elements was that these patients in a highly selected kind of way may be few and far between, the problem of equivocal cases, and that may be a harder to study population, and may be justification for the way I voted.

DR. PAPATHEOFANIS: That makes sense. I think it also points to the use of information that is not in the peer reviewed literature and how that was used in this example, for those who are in attendance in the audience, to really make a decision on a really hard one. And I agree, hopefully it will be a population that's very well screened and preselected and the technology is used appropriately.

So with that, we have one more.

DR. MCNEIL: No, we have two more.

DR. PAPATHEOFANIS: Oh, we have to do the replacement. I apologize.

00271

DR. BURKEN: I would kindly ask the chair to ask the committee to vote on the size of the effect, since we did vote in the affirmative on this question, to examine that question of effect size in keeping with the EC recommendations several months ago.

DR. PAPATHEOFANIS: Well, let's have a discussion on the effect size, if you want to start us off with that, Mitch.

DR. BURKEN: Earlier, to take us back to 8:30 this morning, I had talked about the seven-point scale that was recommended by the Executive Committee, and I would ask that the

panelists consider placing this effect size into a range ranging from not effective to less effective, as effective, more effective, and breakthrough technology, with some breakdowns within the less effective and as effective range, as I talked about this morning.

I know it's not, you know, not the easiest thinking to break down some of this stuff that has some intangibles into a neat discrete category, but I would ask that we give it our best shot.

MS. NOVAK: The way the question is

00272

worded, it says better health outcomes, so I think we're almost voting that it's more effective, or has the potential of adding something.

DR. GUYTON: It could be as effective with advantages.

DR. PAPATHEOFANIS: Right. The degree is what he is getting after.

MS. NOVAK: Yes, but the improved health outcomes to me as part of the definition would be that means it is more effective, because we said it has improved health outcomes.

DR. PAPATHEOFANIS: But I think in that sense it's from the baseline condition of the patient who's having the study done. I think it's not a generic improvement of health outcomes. In other words, someone who's ill who experiences --

DR. GUYTON: That's not the way I interpreted it.

DR. PAPATHEOFANIS: How would you interpret it?

DR. GUYTON: Compared to the other strict staging tests that are available.

DR. PAPATHEOFANIS: Okay, you can do that. We still need an effect size.

DR. GUYTON: Then it's either as

00273

effective with advantages or more effective, as far as I'm concerned. I don't consider it

breakthrough.
DR. PAPATHEOFANIS: We have seven categories?
DR. BURKEN: That's correct.
DR. PAPATHEOFANIS: We're not going to go to one extreme or the other so we won't, and I don't mean to put words in anyone's mouth, but you just suggested it's not a breakthrough. If anyone think it's not effective at all, obviously you wouldn't have voted the way you did, so the two extremes are pretty much out.
DR. BURKEN: Let me clarify that, Dr. Papatheofanis. You can have adequate evidence in part one but the evidence could be extremely negative, at which point it would be not effective. In this case, we've had some evidence that is positive, but you know, so that has kind of taken not effective out of the picture.
DR. PAPATHEOFANIS: So what are the middle five categories, or what categories are missing, I should ask.
DR. BURKEN: Well, the middle five is less effective without any advantages such as

00274

tolerability or convenience, less effective with advantages, as effective without advantages, as effective with advantages, or more effective. And I know these are kind of slippery categories in spots, but again, this is just a framework that was put in front of us several months ago.
DR. PAPATHEOFANIS: Okay. Barbara?
DR. MCNEIL: Frank, I guess I -- do we have to vote on this? Because I'm going to abstain, I don't know how to answer the question, because we have, even if the data were compelling, you know, if it were 70 percent instead of 50.1, the health outcomes are a little bit hard for me to quantify on this scale. I understand what Jeff said is really what it's doing is improving treatment strategies, and the associated health outcomes are going to vary with what treatment is changed to what for what organ.

DR. GUYTON: Isn't the effectiveness of the detection, because we're saying that it's used to detect locoregional recurrence and distant metastases, and the question is, how effective is it in detecting locoregional disease or metastasis or recurrence.

DR. MCNEIL: I don't think that's how

00275

they formulated the --

MR. KLEIN: I think we may want to hinge where we score this on the choice of words adjunctive versus replacement, because I think the reason it was for adjunctive as opposed to replacement has some relationship to the perceived effectiveness. I'm curious what effectiveness with advantages typically means, Mitch.

DR. BURKEN: What I'm going to do is kind of answer your question in a more reflective manner. These are interim guidelines that have been suggested by the Executive Committee. It turns out that when we put them into play, they may not play out as easily as we would have liked, you know, so the very fact that we are wrestling and grappling with this and maybe having a hard time with it may mean that we need to go back to the Executive Committee and consider some other ways of trying to quantity or scale these effects, or maybe not scaling them all. So I would leave it up to the committee to try to wrestle, and the Executive Committee can I think get some good feedback from this discussion.

DR. PAPATHEOFANIS: Donna?

MS. NOVAK: I think the with

00276

advantages, I can see if it's noninvasive, versus the current procedure which is invasive, if there's a quicker diagnostic time because of whatever, I would think that's what you're talking about with advantages, does it have, you know --

There's two things I think you're asking for. One, is it more effective than the

current procedures and I agree with you, that it's effective in diagnosing, not in curing, although there is certainly some type of relationship there. And then the second is, is there any advantage over the current procedures, and I don't think we've heard any testimony about that at all. I mean, we might know, but I don't think there is anything that has been written up as far as advantages.

DR. BURKEN: And again, the two potential types of advantages that come quickly to mind are convenience and tolerability of a particular tests, and maybe others.

DR. PAPATHEOFANIS: The others that are listed, the language in the interim guidelines is convenience, rapidity of effect, fewer side effects, and other advantages, and that's under category three, which is effective but with

00277

advantages.

So, let's say we don't want to vote on size of health effects, or the committee chooses not to. Is that an option?

DR. BURKEN: I certainly think it would be, and it would send certainly some message to the Executive Committee.

MS. NOVAK: Are we going to vote on whether we are going to vote, or is that --

DR. PAPATHEOFANIS: Well, we need a motion. If there is no more discussion, we need a motion that you're not going to vote on it, and so if the motion is that, then we will vote on the fact that you will not vote on this.

DR. GUYTON: I guess my question is is it, are we voting on the effectiveness of the improvement in health outcomes or the effectiveness of the test in detecting disease. That's the real issue, I think.

DR. BURKEN: To go back to this morning when I stood up, I said everything we're talking about in PET today is compared to something else, so it would be the effectiveness of PET versus

conventional diagnostic tests, so that's, you should always think of PET and its comparator or

00278

comparators.

DR. PAPATHEOFANIS: Does that help?

DR. GUYTON: Well, I guess it still raises the question of are we comparing it in terms of detecting disease or improving health outcomes.

DR. BURKEN: Improving health outcomes is what we're trying to do here.

DR. ABRAMS: But, I just think this gets to what I brought up earlier. I mean, the only thing that an imaging modality can really do is help you make a different decision about the treatment, and that eventually depending on how good that treatment is, will or will not affect the overall health outcome. But if you use health outcome in a very broad sense, making a different decision, you know, give radiation, not give radiation, that, I mean, I think if it didn't affect the decision, because we went through this earlier, then you wouldn't want to use it as an adjunct.

I mean, if you were just doing it to have another test, it doesn't make any sense, so I, you know, I take this as a whole, and that's just a subjunctive clause in the sentence, and

00279

basically my thinking about this was the reason it could be as effective or more effective in certain circumstances. This advantages and stuff, I must confess, I'm not sure what that really means in this context. But it has to be as effective or else you wouldn't have voted yes, and it may be more effective in certain circumstances.

DR. LERNER: I'm just wondering whether we're trying to fit a square peg into a round hole. I think that maybe this was developed to be an overall set of categories, and what we really found here was something that doesn't quite fit.

DR. PAPATHEOFANIS: Isn't generalizable?
DR. LERNER: Right. And maybe we're better off rather than trying to make that fit in an uncomfortable way to simply say back to the Executive Committee, maybe for situations like this, you need something else.
DR. GUYTON: Actually, I agree with Jeff with regard to his characterization that, improving health outcomes in that regard, and I will move the question.
DR. PAPATHEOFANIS: What is the question then?

00280

DR. GUYTON: Vote on the effectiveness.
DR. PAPATHEOFANIS: To go ahead and vote on the effectiveness?
DR. GUYTON: Right.
DR. PAPATHEOFANIS: And how would you categorize it in the seven categories?
DR. GUYTON: If you want me to categorize it in the seven, I'll say more effective.
DR. PAPATHEOFANIS: So more effective, the new intervention improves health outcomes by a significant albeit small margin, as compared with established services or medical items.
DR. GUYTON: Uh-huh.
MS. NOVAK: I will second that.
DR. PAPATHEOFANIS: Any discussion on that? Let's go to a vote.
MS. ANDERSON: Okay. We're voting on whether the effect size is considered more effective, just the language more effective. Those voting for? Those voting against? And those abstaining.
I believe we have two votes for, one vote against, and three abstentions. That means the vote does not carry, but the information is in

00281

the record.

DR. PAPATHEOFANIS: Let's go back to question four but with the different language that includes the words replacement for, and I'd just like to move ahead.

MS. NOVAK: Excuse me. I have a problem with leaving the "or" in. I think maybe what you have to do is take out the "adjunct to" and then vote on the other half of the question, the replacement for.

DR. PAPATHEOFANIS: I think that would be clear, so does someone want to provide some language here?

DR. FLAMM: Improves health outcomes as a replacement for, blah, blah, blah.

(Inaudible colloquy.)

MS. ANDERSON: So what we're voting on, is there adequate evidence that PET improves health outcomes as a replacement for standard staging tests in detecting locoregional recurrence or distant metastases recurrence? Those voting for? Those voting against? We have a unanimous against vote.

DR. PAPATHEOFANIS: Great. Let's move on. Last question. Is there adequate evidence

00282

that PET can improve health outcomes by providing either a more accurate or an earlier determination of tumor response to treatment compared to the use of conventional response criteria which may rely upon clinical exam and/or standard imaging tests, for example CT, MR or bone scan. Any discussion on this one?

DR. TUNIS: Let me just maybe mention one thing again from our previous coverage policy related to monitoring therapy, which is, what we cover for the other oncologic indications is restaging after the completion of a planned course of chemotherapy or therapy, but that monitoring during a planned course of treatment to look for tumor response is not covered. That's for the other cancers just so you know what existing policy is, so we did not elect to cover monitoring

a response to therapy during a planned course of treatment, but did allow for coverage of restaging following the completion of a planned course of treatment.

DR. PAPATHEOFANIS: Any discussion? Any additional information that anyone would need from the audience on this one?

DR. WAHL: If I could comment? It

00283

appears not.

DR. PAPATHEOFANIS: Go ahead.

DR. WAHL: Sean mentioned that the response therapy wasn't covered in the other tumor, since you pointed out, it should be mentioned that breast cancer has been studied probably more extensively in terms of sequential studies and response to treatment than many of the other cancers, and probably that's why you are considering it as a fifth question.

DR. TUNIS: Well, I wasn't proposing that that should be the model for this coverage, just that they should know what the coverage was for the other. It seems to me this hinges a lot, and maybe Dr. Abrams, you could fill in here to what extent the treatment of breast cancer is a trial and error or multiple options, you try something and you look for signs of recurrence and how often those signs -- I means signs of response, and how often things that are detected, how PET could add something there. Does that question make sense?

DR. ABRAMS: Yeah. I mean, I think the type of research studies that have been presented the us and that are being published, the Mortimer

00284

study that recently was published, the hormonal therapy study that Dr. Wahl cited, these are very exciting, because if we could have something that we could rely on like a PET scan fairly soon into a treatment to tell us that treatment was working and we didn't have to wait for the longer end

point of response rate on standard scans that usually takes at least a minimum of four weeks and maybe eight weeks, you could spare people treatment that wasn't helping them, and I think that would be beneficial in some cases.

But I, you know, I don't think from my reading of this yet, that the evidence supports that. I think that what it supports is that these studies, again, need to be done and there is at least sufficient evidence to do more of this type of research and that it's promising, but I don't know that I read anything that convinces me that it's ready to be used in lieu of the standard tests at this point.

DR. PAPATHEOFANIS: Thank you. Dr. Flamm.

DR. FLAMM: I agree that some of these studies are interesting and provide some provocative results, but they are small studies

00285

and one concern I have, especially for the studies that do report imperfect prediction of tumor response is that at least if a patient is going to go on and respond to the treatment that they're on, and you because of your PET think that they are a nonresponder and you take them off of that treatment to which they ultimately would have responded and put them onto some second line maybe not as effective treatment regimen, what have you done to that patient, have you really helped them. That's one of my concerns.

DR. PAPATHEOFANIS: Okay. Any other discussion? We need a motion to take a vote then on question number five.

DR. MCNEIL: I move to call the question.

DR. MANYAK: Second.

DR. PAPATHEOFANIS: Okay.

MS. ANDERSON: The question reads as follows: Is there adequate evidence that PET can improve health outcomes by providing either a more accurate or an earlier determination of tumor

response to treatment compared to the use of conventional response criteria, which may rely upon clinical exam and/or standard imaging tests

00286

such as CT, MRI or bone scan.

Those voting for? Those voting against? And no abstentions. That's a unanimous vote against.

DR. PAPATHEOFANIS: Great. That fulfills the charge of this committee. I want to spend five minutes, and I know everyone has a flight, but mine is not until 5:20. Everyone has a flight to catch, but I want to spend five minutes going back and touching on what recommendations, if any, we can make that are specific and that we think might be of use to HCFA as far as the future role of this indication and the use of this technology. Do you want to start, Steve?

DR. GUYTON: I think I went through the potential studies that HCFA might do. I would caution them to try to avoid some of the contentious parts of the NETT trial.

DR. PAPATHEOFANIS: So you're not a big NETT proponent. Anyone else? Jeff.

DR. LERNER: It's not a study per se, but I'm just wondering whether future MCAC would ever want to issue some guidance to people presenting for public comments to MCAC committees

00287

that might, you know, help them know sort of from the get-go what panels tend to look for in information, and I think it might be helpful.

DR. PAPATHEOFANIS: Well, I think that's a good recommendation. I think the folks who spoke from the public sector and from other vantages did an excellent job today, I think the discussions were very focused, I think the comments were relevant. And so they are picking up that sort of guidance from what's out there, but I think it can be refined maybe a step further

to some specifics, I think that's what you're getting at.

DR. LERNER: Right. I want to be clear, I wasn't being critical of what people presented, I just think there's a way to make it easier for them to have a sense of what some of the expectations are as a panel, and it's moving for efficiency, not a criticism.

MS. NOVAK: Along those same lines of process, it would be helpful possibly if you allowed individuals that were going to testify to provide that ahead of time, and gave some time frame of when they would have to get it in in order to get it disseminated before we caught

00288

airplanes.

DR. ABRAMS: Having watched this play out, new treatments, and I am more familiar with drug treatments, but in other areas in breast cancer, it seems until you have some partnership of the research arm, the payer arm and industry, and what the mix should be in any given treatment may be different depending on the financial circumstances, but until you do that, I don't think you will get these trials done. So although we recommended a lot of I think good trials and ones that people would like to do, it is true that if they are not going to find some way to pay for them they will not be done, and we will all be frustrated sitting around asking these questions in another five years.

So I would hope that perhaps payers would see it to their advantage to some degree to pay for patients in trials so that they don't have to support costs outside of trials which don't answer the question and which perpetuates this sort of lack of information.

DR. GUYTON: It wouldn't necessarily have to be limited to payers, it could be manufacturers or pharmaceutical companies or

00289

whatever.

DR. ABRAMS: That's why I said partnership, I think it has to be a partnership.

DR. GUYTON: It may have to be some sort of request for proposal that includes not only the effect that the research will have, but how much it will cost the government.

DR. PAPATHEOFANIS: That's great.

Just to give you a thumbnail of what happens next, our deliberations as I say will be summarized and passed along to Sean and then eventually to the Executive Committee for ratification, so there will be some discussion again at the Executive Committee. The level of discussion, the details and so forth, we really can't predict, but that's what happens next, then ratification, and then I guess up to the Administrator's office is the next step.

Did you have anything to add before I close?

DR. TUNIS: Just as many of you know, the ratification function of the Executive Committee is set to expire as of October 1st of this year. However, the decisions that are made until that goes into effect will probably still be

00290

subject to Executive Committee ratification. As far as I know, we don't have a scheduled EC meeting, or do we have a tentative.

MS. CONRAD: October 17th.

DR. TUNIS: October 17th is the tentative Executive Committee date. If we could find a way to get this finalized prior to then, which is not out of the question, we will certainly pursue that, but I guess Frank and Barbara will be writing up their detailed summary of the deliberations here, which will take a little bit of time, to present to the EC.

DR. PAPATHEOFANIS: Well, I wanted to thank Sean for being here, and to thank Janet Anderson for her efforts in getting this thing together, and all the committee members. There

were rumors that some of us had died in the two-year interval since we met, but hopefully we will meet again sooner than two years.

MS. ANDERSON: Actually, we are not done yet. We have to stay in compliance so there's two more things we have to do.

The first being that I do want to remind everyone that continuing information can be found on our web site. Our name may have changed,

00291

but the web site is the same, www.hcfa.gov/coverage, or you can just go to www.hcfa.gov and click on the coverage process.

Now, to conclude today's session, would someone please move that the meeting be adjourned.

DR. MANYAK: So move.

DR. GUYTON: Second.

MS. ANDERSON: And second, thank you. The meeting is adjourned.

(Whereupon, the meeting adjourned at 3:56 p.m.)